ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

# Στοιχεία Πιθανοτήτων και Στατιστικής

## Γ΄ ΓΕΝΙΚΟΥ ΛΥΚΕΙΟΥ

Ομάδας Προσανατολισμού Ανθρωπιστικών Σπουδών

ΙΝΣΤΙΤΟΥΤΟ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΕΚΔΟΣΕΩΝ «ΔΙΟΦΑΝΤΟΣ»

# Στοιχεία Πιθανοτήτων
# και Στατιστικής

## ΑΝΑΜΟΡΦΩΣΗ, ΠΡΟΣΑΡΜΟΓΗ ΚΑΙ ΕΠΙΜΕΛΕΙΑ ΠΕΡΙΕΧΟΜΕΝΟΥ

**Μιχαήλ Λουλάκης,** Αναπληρωτής Καθηγητής Ε.Μ.Π.
**Διαμαντίδης Δημήτριος,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03)
**Στουραΐτης Κωνσταντίνος,** Σύμβουλος Α΄ ΙΕΠ

Το παρόν εκπονήθηκε αμισθί, με ευθύνη του ΙΕΠ με βάση την υπ' αριθμ. 27/29-05-2020 Πράξη του ΔΣ του ΙΕΠ

**ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ**

**Ιωάννης Αντωνίου**

Πρόεδρος του Ινστιτούτου Εκπαιδευτικής Πολιτικής

## ΑΡΧΙΚΗ ΕΚΠΟΝΗΣΗ

| | |
|---|---|
| **ΣΥΝΤΟΝΙΣΜΟΣ:** | **Σκούρας Αθανάσιος,** Σύμβουλος Α΄ ΥΠ.Π.Ε.Θ |
| **ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΟΜΑΔΑ ΜΕΛΩΝ ΔΕΠ:** | **Θεοδόσιος Ζαχαριάδης,** Καθηγητής Ε.Κ.Π.Α.<br>**Μιχαήλ Λουλάκης,** Αναπληρωτής Καθηγητής Ε.Μ.Π.<br>**Δέσποινα Πόταρη,** Καθηγήτρια Ε.Κ.Π.Α. |
| **ΑΝΑΜΟΡΦΩΣΗ ΚΑΙ ΕΠΙΜΕΛΕΙΑ ΠΕΡΙΕΧΟΜΕΝΟΥ:** | **Σκούρας Αθανάσιος,** Σύμβουλος Α΄ ΥΠ.Π.Ε.Θ.<br>**Στράντζαλος Αθανάσιος,** Σύμβουλος Β΄ Ι.Ε.Π. |
| **ΕΚΠΟΝΗΣΗ:** | |
| *Εξωτερικοί εμπειρογνώμονες:* | **Βρούτσης Νικόλαος,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03)<br>**Διαμαντίδης Δημήτριος,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03)<br>**Καρκάνης Βασίλειος,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03)<br>**Μηλιώνης Χρίστος,** Σχολικός Σύμβουλος (ΠΕ03)<br>**Πατσαλιάς Μιχάλης,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03)<br>**Σκουρκέας Αναστάσιος,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03)<br>**Σπαθάρας Δημήτριος,** Σχολικός Σύμβουλος (ΠΕ03)<br>**Στουραΐτης Κωνσταντίνος,** Εκπαιδευτικός Δημοσίου Τομέα (ΠΕ03) |
| *Μέλη της Μονάδας Φυσικών Επιστημών, Τεχνολογίας, Μαθηματικών του ΙΕΠ* | **Σκούρας Αθανάσιος,** Σύμβουλος Α΄ ΥΠ.Π.Ε.Θ.<br>**Στράντζαλος Αθανάσιος,** Σύμβουλος Β΄ Ι.Ε.Π. |

Δράση για την αναμόρφωση ή/και εκπόνηση Προγραμμάτων Σπουδών και τυχόν συμπληρωματικού εκπαιδευτικού υλικού στα θεματικά πεδία των Ανθρωπιστικών Επιστημών, Κοινωνικών Επιστημών, Φυσικών Επιστημών και Μαθηματικών Πρωτοβάθμιας και Δευτεροβάθμιας Εκπαίδευσης (Πράξη 36/1 4-09-201 7 του ΔΣ του ΙΕΠ και σε συνέχεια την με αρ. πρωτ. 6139/1 9-09-201 7 και ΑΔΑ: 73Μ1 ΟΞΛΔ-ΗΨΟ Πρόσκληση εκδήλωσης ενδιαφέροντος)

**ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ**
**Γεράσιμος Κουζέλης**
Πρόεδρος του Ινστιτούτου Εκπαιδευτικής Πολιτικής

**Υπεύθυνη Δράσης**
**Γεωργία Φέρμελη**
Σύμβουλος Α΄ του Ινστιτούτου Εκπαιδευτικής Πολιτικής

Το παρόν εκπονήθηκε αμισθί, με ευθύνη της Μονάδας Φυσικών Επιστημών, Τεχνολογίας και Μαθηματικών του Ινστιτούτου Εκπαιδευτικής Πολιτικής, στο πλαίσιο της ανωτέρω Δράσης.

**ΠΡΟΕΚΤΥΠΩΤΙΚΕΣ ΕΡΓΑΣΙΕΣ: Ι.Τ.Υ.Ε. "ΔΙΟΦΑΝΤΟΣ"**

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΙΝΣΤΙΤΟΥΤΟ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

# Στοιχεία Πιθανοτήτων και Στατιστικής

## Γ΄ Λυκείου

Ομάδας Προσανατολισμού
Ανθρωπιστικών Σπουδών

ΙΝΣΤΙΤΟΥΤΟ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΕΚΔΟΣΕΩΝ «ΔΙΟΦΑΝΤΟΣ»

# ΠΡΟΛΟΓΟΣ

Με το παρόν εκπαιδευτικό υλικό επιχειρείται η «μεταφορά» των περιεχομένων , των εννοιών και ιδεών του Προγράμματος Σπουδών των Μαθηματικών Γενικής Παιδείας της Γ΄ Λυκείου στην καθημερινή διδακτική πρακτική. Ακολουθώντας τους σκοπούς και τους στόχους του ΠΣ, περιγράφεται συνολικά μία πρόταση-πορεία (ένα «παράδειγμα») εφαρμογής του νέου ΠΣ. Για τον λόγο αυτό κρίθηκε σκόπιμο όπως το παρόν υλικό προσομοιάζει με διδακτικό εγχειρίδιο ώστε ο μαθητής να το χρησιμοποιεί ως βασική πηγή για τη οργάνωση της μελέτης του.

Το υλικό κρίθηκε επίσης σκόπιμο, να έχει κατά το δυνατόν κοινή διάρθρωση σε όλες τις παραγράφους του. Για αυτό τον λόγο αποτελείται από τα παρακάτω διακριτά τμήματα, προσπαθώντας να περιγράψει την πορεία διδασκαλίας – μάθησης που εξελίσσεται σε μια σχολική τάξη:

- Διερεύνηση,
- Βασικές μαθηματικές έννοιες - ιδέες - διεργασίες,
- Εφαρμογές,
- Ασκήσεις - προβλήματα - δραστηριότητες,
- Πρόσθετο υλικό.

Το περιεχόμενο έγινε προσπάθεια να είναι επαρκές, ώστε να υποστηρίζει τη μάθηση των Μαθηματικών ενεργοποιώντας όλους τους μαθητές και αναδεικνύοντας τη σημασία τους, ως μάθημα Γενικής Παιδείας, απαραίτητο για τον αυριανό πολίτη, ανεξάρτητα από τον επαγγελματικό του προσανατολισμό και τα επιστημονικά του ενδιαφέροντα. Ασφαλώς και υπάρχουν περιθώρια βελτίωσης και ιδίως από τις παρατηρήσεις των διδασκόντων, που μέσα από την πραγματικότητα της σχολικής τάξης λειτουργούν αυθεντικά ως διαμεσολαβητές των αναγκών των μαθητών τους και παρέχουν ανατροφοδότηση, ώστε το εκπαιδευτικό υλικό να παραμένει πάντα «ζωντανό» και επίκαιρο.

# ΠΕΡΙΕΧΟΜΕΝΑ

# ΠΙΘΑΝΟΤΗΤΕΣ

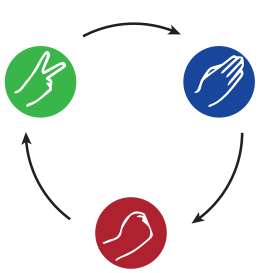
*Εισαγωγή*

Η Θεωρία των Πιθανοτήτων είναι ένας σύγχρονος κλάδος των Μαθηματικών που ξεκίνησε να αναπτύσσεται κυρίως μετά τον 16ο αιώνα, αλλά ακόμα περισσότερο τον 18ο αιώνα με τις εργασίες των διάσημων μαθηματικών Bernoulli, de Moivre, Laplace και Gauss. Είναι διαδεδομένη η άποψη ότι η Θεωρία Πιθανοτήτων έχει κυρίως ως εφαρμογή την μοντελοποίηση παιχνιδιών. Στην πραγματικότητα είναι πολύ περισσότερες οι εφαρμογές της στην Πληροφορική, τις Τηλεπικοινωνίες, την Ιατρική, τη Φυσική, σε Θετικές, Κοινωνικές και άλλες επιστήμες. Ήδη από τον 18ο αιώνα ο Laplace δεν περιορίζεται μόνο στη μαθηματική ανάλυση παιχνιδιών, αλλά εφαρμόζει τα συμπεράσματά του και σε ένα πλήθος από επιστημονικά και πρακτικά προβλήματα. Στις μέρες μας η Θεωρία των Πιθανοτήτων με εργασίες διάσημων μαθηματικών όπως οι Chebyshev, Markov, Von Mises και ιδιαίτερα με την αξιωματική θεμελίωση των Πιθανοτήτων από τον Kolmogorov 1933, έχει σημειώσει αλματώδη πρόοδο, παρέχοντας χρήσιμα μοντέλα και μεθόδους ανάλυσης για πολύπλοκα προβλήματα.

Σε αυτό το κεφάλαιο θα μελετήσουμε τα θεμελιώδη στοιχεία της Θεωρίας Πιθανοτήτων και θα εμπλακούμε στη μοντελοποίηση πολύπλοκων φαινομένων, στα οποία υπάρχει αβεβαιότητα ως προς την έκβασή τους.

## ΕΝΟΤΗΤΑ 1.1. ΠΕΙΡΑΜΑΤΑ ΤΥΧΗΣ, ΔΕΙΓΜΑΤΙΚΟΣ ΧΩΡΟΣ ΚΑΙ ΕΝΔΕΧΟΜΕΝΑ

### Διερεύνηση



Η Άννα και ο Βασίλης παίζουν το γνωστό παιχνίδι «πέτρα, ψαλίδι, χαρτί».

α) Ο Βασίλης σκέφτηκε να διαλέγει συνέχεια «πέτρα». Πιστεύετε ότι αυτή η στρατηγική του δίνει πλεονέκτημα έναντι της Άννας;

β) Η Άννα πρότεινε να αλλάξουν το παιχνίδι και να παίξουν «πέτρα, ψαλίδι, μολύβι, χαρτί». Αν ο Βασίλης διατηρήσει την ίδια στρατηγική τότε ευνοείται από την αλλαγή του παιχνιδιού ή όχι, κατά τη γνώμη σας;

## Βασικές μαθηματικές έννοιες-Ιδέες-Διεργασίες

***Πειράματα τύχης***

Στο περιβάλλον μας υπάρχουν φαινόμενα που μπορούμε να προβλέψουμε την εξέλιξή τους, αν γνωρίζουμε τις συνθήκες υπό τις οποίες συμβαίνουν. Για παράδειγμα, αν έχουμε ένα δοχείο με αποσταγμένο νερό σε θερμοκρασία $32^o$ Κελσίου και το θερμάνουμε μέχρι η θερμοκρασία του να φτάσει τους $100^o$ Κελσίου, τότε το νερό θα βράσει. Κάθε τέτοιο πείραμα, που η γνώση των συνθηκών κάτω από τις οποίες εκτελείται μας επιτρέπει να προκαθορίσουμε πλήρως το αποτέλεσμά του, λέγεται αιτιοκρατικό (deterministic) πείραμα.

Από την άλλη μεριά, δεν μπορούμε να προβλέψουμε την εξέλιξη φαινομένων, όπως το πλήθος των κλήσεων ή των δεδομένων που θα δεχτεί ένα τηλεπικοινωνιακό κέντρο μέσα σε μία ώρα ή τη διάρκεια ζωής μιας λάμπας κτλ. Τέτοια πειράματα, το αποτέλεσμα των οποίων δεν μπορούμε εκ των προτέρων να προβλέψουμε και υπάρχει αβεβαιότητα ως προς την έκβασή τους, μολονότι επαναλαμβάνονται (φαινομενικά τουλάχιστον) κάτω από τις ίδιες συνθήκες, ονομάζονται πειράματα τύχης (random experiments) και μελετώνται από τη Θεωρία των Πιθανοτήτων.

Υπάρχουν πειράματα που έχουν «φύση» αιτιοκρατικού πειράματος, αλλά η αβεβαιότητα της έκβασής τους είναι πολύ μεγάλη, λόγω της πολυπλοκότητας του αιτιοκρατικού μοντέλου που τα περιγράφει. Ένα παράδειγμα είναι το στρίψιμο ενός αμερόληπτου (τίμιου) κέρματος: Αν κανείς γνωρίζει με ακρίβεια τις συνθήκες του στριψίματος, θεωρητικά είναι σε θέση να προβλέψει την έκβαση του πειράματος (δηλ. κεφαλή ή γράμματα) με βάση τους νόμους της κλασικής Φυσικής. Ωστόσο, το αποτέλεσμα του στριψίματος είναι εξαιρετικά «ευαίσθητο» σε πολύ μικρές αλλαγές των αρχικών συνθηκών και έτσι η εξαγωγή ενός τύπου που θα προέβλεπε την έκβαση του στριψίματος από τις αρχικές συνθήκες του πειράματος θα ήταν όχι μόνο εξαιρετικά δύσκολη στην πράξη αλλά και μη λειτουργική, λόγω της πολύπλοκης εξάρτησης του αποτελέσματος από τις αρχικές συνθήκες. Σε αυτή την περίπτωση, είναι πιο αποτελεσματικό να χρησιμοποιηθεί η Θεωρία Πιθανοτήτων, ώστε να μοντελοποιηθεί το στρίψιμο του κέρματος βάσει των πιθανών εκβάσεών του, δηλαδή ως πείραμα τύχης. Με άλλα λόγια, θα λέγαμε ότι «η τυχαιότητα μοντελοποιεί την πολυπλοκότητα».

Παιχνίδια όπως το στρίψιμο ενός κέρματος, η ρίψη ενός ζαριού, το παιχνίδι «πέτρα, ψαλίδι, χαρτί» ή το τάβλι είναι χαρακτηριστικά παραδείγματα πειραμάτων τύχης. Ωστόσο, η Θεωρία των Πιθανοτήτων, εκτός από διασκεδαστικά παιχνίδια, χρησιμοποιείται για την μοντελοποίηση και την ανάλυση προβλημάτων από τον χώρο των Φυσικών Επιστημών, των Επιστημών Υγείας, της Οικονομικής Επιστήμης και άλλων, τα οποία είναι αρκετά πολύπλοκα για να διερευνηθούν ως αιτιοκρατικά πειράματα.

***Δειγματικός χώρος***

Αν $\omega_1$, $\omega_2$,...,$\omega_\kappa$ είναι τα δυνατά αποτελέσματα ενός πειράματος τύχης, τότε ο δειγματικός χώρος του πειράματος θα είναι το σύνολο:
$\Omega = \{\omega_1, \omega_2, \ldots, \omega_\kappa\}$.

Όλα τα αποτελέσματα που μπορούν να εμφανιστούν σε μία εκτέλεση ένος πειράματος τύχης λέγονται δυνατά αποτελέσματα ή δυνατές εκβάσεις του πειράματος. Για παράδειγμα, αν θεωρήσουμε ως πείραμα τύχης την καταγραφή των επιλογών των δύο παικτών του παιχνιδιού «πέτρα, ψαλίδι, χαρτί», τότε ένα δυνατό αποτέλεσμα είναι το (πέτρα, χαρτί), όπου ο πρώτος παίκτης επιλέγει «πέτρα» και ο δεύτερος «χαρτί». Το σύνολο των δυνατών αποτελεσμάτων λέγεται δειγματικός χώρος (sample space, εν συντομία δ.χ.) και συμβολίζεται συνήθως με το γράμμα Ω. Στο συγκεκριμένο παράδειγμα, μπορούμε να θεωρήσουμε ότι ο δ.χ. αποτελείται από όλα τα δυνατά διατεταγμένα ζεύγη των λέξεων «πέτρα», «ψαλίδι» και «χαρτί». Δηλαδή είναι:

$$\Omega = \begin{Bmatrix} (\text{πέτρα}, \text{χαρτί}), (\text{χαρτί}, \text{πέτρα}), (\text{πέτρα}, \text{ψαλίδι}), \\ (\text{ψαλίδι}, \text{πέτρα}), (\text{χαρτί}, \text{ψαλίδι}), (\text{ψαλίδι}, \text{χαρτί}) \end{Bmatrix}$$

Αν θεωρήσουμε ως πείραμα τύχης τη ρίψη ενός ζαριού και μας ενδιαφέρει η ένδειξη της άνω έδρας, τότε μπορούμε να θεωρήσουμε ότι ο δ.χ. είναι:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

***Ενδεχόμενα***

Οποιοδήποτε σύνολο δυνατών αποτελεσμάτων του πειράματος τύχης ονομάζεται ενδεχόμενο (event) ή γεγονός. Για παράδειγμα, στη ρίψη ενός ζαριού κάποια ενδεχόμενα είναι τα παρακάτω Α, Β και Γ:

Α: «Έρχεται άρτιος αριθμός»,

Β: «Έρχεται αριθμός μικρότερος του 5»,

Γ: «Έρχεται 1».

Αν το αποτέλεσμα ενός πειράματος τύχης είναι το $\omega_\nu$ τότε όλα τα ενδεχόμενα που περιέχουν το $\omega_\nu$ λέμε ότι πραγματοποιούνται, ενώ τα ενδεχόμενα που δεν περιέχουν το $\omega_\nu$ λέμε ότι δεν πραγματοποιούνται.

Κάθε ενδεχόμενο αντιστοιχεί σε ένα σύνολο στοιχείων του δ.χ.

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

δηλαδή σε ένα υποσύνολο του Ω. Τα παραπάνω ενδεχόμενα γράφονται :

$$A=\{2, 4, 6\},\ B=\{1, 2, 3, 4\} \text{ και } \Gamma=\{1\}$$

Αν γίνει μία ρίψη του ζαριού και το αποτέλεσμα είναι 3, τότε λέμε ότι το Β πραγματοποιείται γιατί περιέχει το 3, ενώ τα Α και Γ δεν πραγματοποιούνται, γιατί δεν περιέχουν το 3. Το 3 είναι ένα ευνοϊκό αποτέλεσμα για το Β. Άλλα ευνοϊκά αποτελέσματα για το Β είναι τα 1, 2 και 4. Αντίστοιχα, ευνοϊκά αποτελέσματα για το Α είναι τα 2, 4 και 6, ενώ για το Γ είναι μόνο το 1.

***Η γλώσσα των συνόλων***

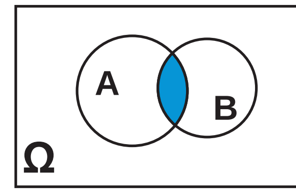
Όπως είδαμε τα ενδεχόμενα είναι υποσύνολα του δ.χ. Ω. Συνεπώς, μπορούν να αναπαρασταθούν με διαγράμματα Venn.

Η τομή των ενδεχομένων Α και Β συμβολίζεται με $A \cap B$ και πραγματοποιείται όταν το Α και το Β πραγματοποιούνται. Για το παραπάνω παράδειγμα είναι:

$$A \cap B = \{ 2, 4 \}$$

Άρα τα στοιχεία της τομής των Α και Β είναι τα κοινά ευνοϊκά αποτελέσματα των δύο ενδεχομένων Α και Β.

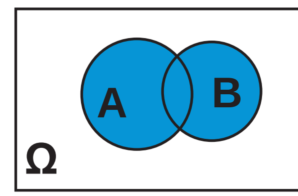
$A \cap B$




Αντίστοιχα, η ένωση των Α και Β συμβολίζεται με $A \cup B$ και πραγματοποιείται όταν ένα τουλάχιστον από τα Α ή Β πραγματοποιείται. Για το παραπάνω παράδειγμα είναι:

$$A \cup B = \{ 1, 2, 3, 4, 6 \}$$

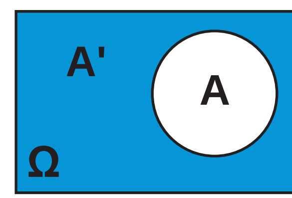
$A \cup B$




Το συμπληρωματικό ή αντίθετο του ενδεχομένου Α συμβολίζεται με Α΄ και πραγματοποιείται, αν το Α δεν πραγματοποιείται. Για το παραπάνω παράδειγμα είναι:

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

$A'$

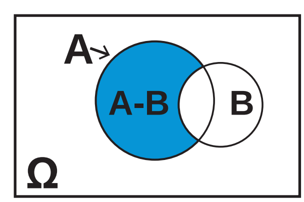



Δηλαδή το Α΄ πραγματοποιείται αν έρθει περιττός αριθμός.

Η διαφορά του Β από το Α συμβολίζεται με Α - Β και πραγματοποιείται όταν το Α πραγματοποιείται αλλά όχι το Β. Η Α – Β αποτελείται από τα ευνοϊκά αποτελέσματα του Α που δεν είναι ευνοϊκά για το Β. Για το παραπάνω παράδειγμα είναι:

$$A - B = \{ 6 \}$$

$A - B$




Δηλαδή, από τα 2, 4 και 6 που είναι ευνοϊκά για το Α «διαγράφουμε» τα 2 και 4 που είναι ευνοϊκά και για το Β. Με άλλα λόγια το Α - Β πραγματοποιείται αν έρθει άρτιος αριθμός που δεν είναι μικρότερος του 5. Ανάλογα, στο παράδειγμά μας $B - A = \{ 1, 3 \}$.

Με αφορμή το παράδειγμα του ζαριού κάνουμε τις παρακάτω γενικές παρατηρήσεις:

- Ένα ενδεχόμενο που περιέχει ένα μόνο στοιχείο του δ.χ., όπως το Γ στο παράδειγμά μας, ονομάζεται απλό ή στοιχειώδες.
- Ενδεχόμενα που περιέχουν περισσότερα από ένα στοιχεία του δ.χ., όπως τα Α και Β στο παράδειγμά μας, ονομάζονται σύνθετα.
- Το ενδεχόμενο $\varnothing$ ονομάζεται αδύνατο ενδεχόμενο.
- Αν δύο ενδεχόμενα δεν έχουν κοινά ευνοϊκά αποτελέσματα ή κοινά στοιχεία ως σύνολα (δηλαδή δεν μπορούν να πραγματοποιηθούν ταυτόχρονα), όπως τα Α και Γ στο παράδειγμά μας όπου $A \cap \Gamma = \varnothing$, λέμε ότι είναι ασυμβίβαστα. Έτσι, τα Α και Γ είναι ασυμβίβαστα.

Το κενό σύνολο $\{ \ \}$ συμβολίζεται ως $\varnothing$.

— Το ενδεχόμενο Ω ονομάζεται βέβαιο ενδεχόμενο.
— Αν η ένωση δύο ενδεχομένων είναι όλος ο δ.χ. Ω, όπως η ένωση του Β και του Δ = { 3, 5, 6 } στο παράδειγμά μας, όποιο κι αν είναι το αποτέλεσμα του πειράματος τύχης, θα πραγματοποιηθεί τουλάχιστον ένα από αυτά: είτε το Δ είτε το Β, είτε και τα δύο.
— Αν ένα στοιχείο Γ είναι υποσύνολο ενός ενδεχομένου Β, όπως στο παράδειγμά μας, όπου $Γ \subseteq B$, αν το Γ πραγματοποιείται, τότε και το Β πραγματοποιείται.

## Εφαρμογές

### Εφαρμογή 1

Η Αθηνά και ο Κωστής προσπαθούν να γράψουν έναν δ.χ. για το πείραμα τύχης των τριών διαδοχικών ρίψεων ενός νομίσματος.

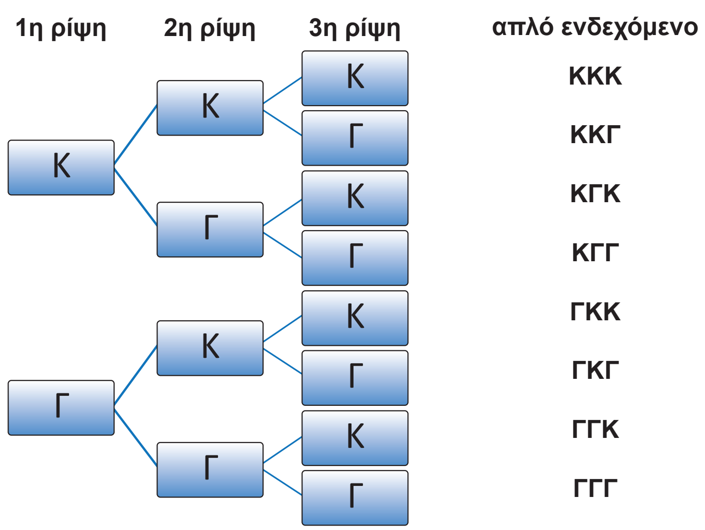
Η Αθηνά γράφει ότι υπάρχουν τα παρακάτω απλά ενδεχόμενα, όπου Κ σημαίνει κεφαλή και Γ σημαίνει γράμματα:

$\omega_1$: 1Κ και 2 Γ, $\omega_2$: 1Γ και 2 Κ, $\omega_3$: 3Κ και $\omega_4$: 3 Γ.

Άρα ο δ.χ. σύμφωνα με την Αθηνά είναι:

$$Ω = \{ \omega_1, \omega_2, \omega_3, \omega_4 \}.$$

Ο Κωστής έδωσε άλλη απάντηση κατασκευάζοντας το παρακάτω δενδροδιάγραμμα:




Άρα ο δ.χ. είναι:

$$Ω = \{ ΚΚΚ, ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΚΚ, ΓΚΓ, ΓΓΚ, ΓΓΓ \}$$

Με ποιον από τους δύο συμφωνείτε;

**Λύση**

Και οι δύο δειγματικοί χώροι είναι σωστοί εφόσον περιέχουν όλες τις πιθανές εκβάσεις του πειράματος τύχης.
Ο δ.χ. του Κωστή περιγράφει με μεγαλύτερη λεπτομέρεια την έκβαση του πειράματος και είναι καταλληλότερος για την μοντελοποίηση του πειράματος τύχης σε περιπτώσεις που είναι σημαντική η σειρά που εμφανίζονται τα Κ και Γ.
Για παράδειγμα, σκεφτείτε το παρακάτω πρόβλημα (1ο σενάριο):
«Σε ένα παιχνίδι με τρεις παίκτες, πρώτος στρίβει το κέρμα ο παίκτης Α, δεύτερος ο Β και τρίτος ο C και κερδίζουν δώρο όποιοι φέρουν Κ».
Αν έχουμε την πληροφορία ότι η έκβαση του πειράματος είναι το $\omega_2$ του δ.χ. της Αθηνάς, δηλαδή ότι 1 παίκτης έφερε Γ και 2 παίκτες έφεραν Κ, τότε δεν μπορούμε να καταλάβουμε ποιοι κέρδισαν, γιατί δε γνωρίζουμε ποιοι είναι οι δύο που έφεραν Κ.
Αν όμως, σύμφωνα με τον δ.χ. του Κωστή, έχουμε την πληροφορία ότι η έκβαση είναι ΚΚΓ τότε κερδίζουν οι Α και Β. Αντίστοιχα, αν η έκβαση είναι ΚΓΚ κερδίζουν οι Α και C και η έκβαση είναι ΓΚΚ κερδίζουν οι Β και C.
Άρα, ο δ.χ. της Αθηνάς δεν είναι κατάλληλος για αυτό το πρόβλημα, γιατί δε «λαμβάνει υπόψη του» τη σειρά εμφάνισης των Κ και Γ.
Τα πράγματα αλλάζουν αν έχουμε το πρόβλημα (2ο σενάριο):
«Ένας παίκτης στρίβει τρεις φορές ένα αμερόληπτο κέρμα και κερδίζει δώρο αν φέρει περισσότερες φορές Κ από Γ».
Τότε και οι δύο δ.χ. είναι κατάλληλοι για τη μοντελοποίηση του πειράματος τύχης και μάλιστα ο δ.χ. της Αθηνάς είναι απλούστερος αφού περιέχει λιγότερα στοιχεία.

Το συμπέρασμα από αυτό το παράδειγμα είναι ότι ο δ.χ. ενός πειράματος τύχης δεν είναι μονοσήμαντα ορισμένος (π.χ. στο 2ο σενάριο και οι δύο δ.χ. είναι κατάλληλοι) και η επιλογή του δ.χ. είναι μέρος της μοντελοποίησης του πειράματος τύχης (π.χ. στο 1ο σενάριο ο δ.χ. της Αθηνάς δεν είναι ένα καλό μοντέλο για το πείραμα τύχης).

## Εφαρμογή 2

Ρίχνουμε δύο ζάρια και καταγράφουμε το αποτέλεσμα της ρίψης. Ω είναι ο δ.χ. του πειράματος τύχης που περιγράφεται στον διπλανό πίνακα.

| | | | | | |
|---|---|---|---|---|---|
| 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 |

Να γράψετε τα παρακάτω ενδεχόμενα με αναγραφή των στοιχείων τους:

Α: «Το αποτέλεσμα και των δύο ρίψεων είναι άρτιος αριθμός»,

Β: «Το αποτέλεσμα των ρίψεων έχει άθροισμα 7»,

Γ: «Το αποτέλεσμα και των δύο ρίψεων είναι μεγαλύτερο του 3».

Στη συνέχεια να γράψετε τα ενδεχόμενα $A'$, $A \cup \Gamma$, $A - \Gamma$ και $B - \Gamma$ με αναγραφή των στοιχείων τους.

| | | | | | |
|---|---|---|---|---|---|
| 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 |

Με πορτοκαλί το Α

**Λύση**

Έχοντας επιλέξει τον παραπάνω δ.χ. για το πείραμα τύχης, διακρίνουμε τα ζάρια σε πρώτο και δεύτερο. Για παράδειγμα, το (1,2) αντιστοιχεί στο αποτέλεσμα που το πρώτο ζάρι φέρνει 1 και το δεύτερο 2, ενώ το (2,1) αντιστοιχεί στο αποτέλεσμα που το πρώτο ζάρι φέρνει 2 και το δεύτερο 1. Κάθε διατεταγμένο ζεύγος σε κελί του πίνακα αντιστοιχεί σε ένα δυνατό αποτέλεσμα της ρίψης των δύο ζαριών.
Από τον πίνακα προκύπτει ότι:

$$A = \{(2,2), (2,4), (2,6), (4,2), (4,4), (4,6), (6,2), (6,4), (6,6)\}$$

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

και

$$\Gamma = \{(4,4), (4,5), (4,6), (5,4), (5,5), (5,6), (6,4), (6,5), (6,6)\}$$

| | | | | | |
|---|---|---|---|---|---|
| 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 |

Με πράσινο το Β

Το Α΄ πραγματοποιείται αν δεν πραγματοποιείται το Α. Το ενδεχόμενο Α΄ είναι το:

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

| | | | | | |
|---|---|---|---|---|---|
| 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 |

Με κίτρινο το Γ

Το $A \cup \Gamma$ πραγματοποιείται αν πραγματοποιείται ή το Α ή το Γ.

Επομένως:

$$A \cup \Gamma = \left\{\begin{array}{l} (2,2), (2,4), (2,6), (4,2), (4,4), (4,5), (4,6), \\ (5,4), (5,5), (5,6), (6,2), (6,4), (6,5), (6,6) \end{array}\right\}$$

| | | | | | |
|---|---|---|---|---|---|
| 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~~ |

Α – Γ: «Διαγράφουμε» τα στοιχεία του Α που ανήκουν και στο Γ

Το Α - Γ πραγματοποιείται αν πραγματοποιείται το Α, αλλά όχι το Γ.

Άρα:

$$A - \Gamma = \{(2,2), (2,4), (2,6), (4,2), (6,2)\}$$

Το Β - Γ πραγματοποιείται αν πραγματοποιείται το Β, αλλά όχι το Γ.
Το Β και το Γ δεν έχουν κοινά στοιχεία, άρα:

$$B - \Gamma = B$$

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Τα χρώματα μιας ομάδας βόλεϊ είναι λευκό, γαλάζιο και μαύρο. Για κάθε παίκτη/παίκτρια η ομάδα δίνει τα εξής ρούχα:

— Τρεις μονόχρωμες μπλουζες: Μία λευκή (Λ), μία γαλάζια (Γ) και μία μαύρη (Μ).

— Τρία μονόχρωμα σορτσάκια, στα ίδια χρώματα με τις μπλούζες.

— Δύο ζευγάρια κάλτσες, ένα μαύρο κι ένα λευκό.

Επιλέγουμε τυχαία μία μπλούζα, ένα σορτσάκι κι ένα ζευγάρι κάλτσες.

**α)** Να γράψετε έναν δ.χ. του πειράματος τύχης.

**β)** Χρησιμοποιώντας τον παραπάνω δ.χ. να βρείτε το ενδεχόμενο Α: «τα ρούχα που επιλέξαμε έχουν ίδιο χρώμα».

**2)** Ρίχνουμε ένα ζάρι δύο φορές. Στον δ.χ. της εφαρμογής 2 να βρείτε τα ενδεχόμενα:

**α)** Α: Το αποτέλεσμα της 1ης ρίψης είναι μεγαλύτερο από το αποτέλεσμα της 2ης ρίψης.

**β)** Β: Το άθροισμα των ενδείξεων των δύο ρίψεων είναι άρτιος αριθμός.

**γ)** Γ: Το αποτέλεσμα της 1ης ρίψης είναι 6.

**δ)** $A \cap B$, $A \cup B$, $B - \Gamma$, $A - \Gamma$ και $\Gamma - A$.

**3)** Ένα κουτί έχει τρεις μπάλες, μία άσπρη, μία κόκκινη και μία μαύρη. Παίρνουμε από το κουτί μια μπάλα τυχαία και καταγράφουμε το χρώμα της. Μετά ξανατοποθετούμε τη μπάλα στο κουτί και επαναλαμβάνουμε άλλη μία φορά την τυχαία επιλογή μπάλας. Έτσι, στο τέλος έχουμε καταγράψει δύο χρώματα (ίδια ή διαφορετικά), ένα για κάθε μπάλα που επιλέξαμε. Να γράψετε έναν δ.χ. για το πείραμα τύχης και στη συνέχεια να απαντήσετε στα ερωτήματα:

**α)** Ποιο είναι το ενδεχόμενο Α: «η πρώτη μπάλα είναι κόκκινη»;

**β)** Ποιο είναι το ενδεχόμενο Β: «η δεύτερη μπάλα είναι κόκκινη»;

**γ)** Να εκφράσετε λεκτικά το ενδεχόμενο $A \cap B$ και να το βρείτε.

**δ)** Να εκφράσετε λεκτικά το ενδεχόμενο $A - B$ και να το βρείτε.

**4)** Να λύσετε την άσκηση 3, αν αυτή τη φορά η μπάλα που εξάγεται την πρώτη φορά δεν επανατοποθετείται στο κουτί πριν τη δεύτερη εξαγωγή μπάλας.

**5)** Ρίχνουμε δύο ζάρια και καταγράφουμε το αποτέλεσμα της ρίψης. Να εκφράσετε λεκτικά τα ενδεχόμενα $A'$, $A \cup \Gamma$, $A - \Gamma$ και $B - \Gamma$, όπου:

Α: «Το αποτέλεσμα και των δύο ρίψεων είναι άρτιος αριθμός»,

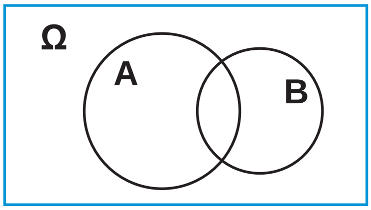
Β: «Το αποτέλεσμα των ρίψεων έχει άθροισμα 7»,

Γ: «Το αποτέλεσμα και των δύο ρίψεων είναι μεγαλύτερο του 3».

**6)** Δύο παίκτες παίζουν σκάκι και συμφωνούν να είναι νικητής εκείνος που πρώτος θα κερδίσει δύο παρτίδες. Να γράψετε έναν δ.χ. για το πείραμα τύχης, από τον οποίο να προκύπτει πόσα παιχνίδια έγιναν μέχρι να βγει νικητής, ποιος προηγήθηκε και ποιος τελικά κέρδισε.

## Πρόσθετο Υλικό

**1)** Οι ένοικοι μίας πολυκατοικίας παρκάρουν τα οχήματά τους σε ένα χώρο στάθμευσης αυτοκινήτων. Στο παρακάτω διάγραμμα Venn, το A έχει ως στοιχεία του ενοίκους που έχουν αυτοκίνητο και το B εκείνους που έχουν μηχανή. Επιλέγουμε τυχαία έναν ένοικο.




Χρησιμοποιώντας τη γλώσσα των συνόλων (τομή, ένωση κτλ.) να εκφράσετε τα ενδεχόμενα ο ένοικος που επιλέγουμε:

**α)** έχει αυτοκίνητο και μηχανή.
**β)** έχει αυτοκίνητο ή μηχανή.
**γ)** δεν έχει αυτοκίνητο ή μηχανή.
**δ)** δεν έχει αυτοκίνητο και δεν έχει μηχανή.
**ε)** δεν έχει (και) αυτοκίνητο και μηχανή.
**στ)** δεν έχει αυτοκίνητο ή δεν έχει μηχανή.
**ζ)** έχει μόνο αυτοκίνητο ή έχει μόνο μηχανή.
**η)** έχει αυτοκίνητο ή μηχανή, αλλά δεν ανήκει σε αυτούς που έχουν και αυτοκίνητο και μηχανή.

**2)** Για δύο ενδεχόμενα A και B ενός δ.χ. Ω να αντιστοιχίσετε τα ενδεχόμενα της πρώτης στήλης με τα ίσα προς αυτά ενδεχόμενα της δεύτερης στήλης.

| 1η στήλη |
|---|
| $(A \cup B)'$ |
| $(A \cap B)'$ |
| $(A - B) \cup (B - A)$ |

| 2η στήλη |
|---|
| $A' \cup B'$ |
| $(A \cup B) - (A \cap B)$ |
| $A' \cap B'$ |

Στη συνέχεια να εκφράσετε λεκτικά τα ενδεχόμενα του πίνακα. Ποιες λεκτικές εκφράσεις αντιστοιχούν σε ίσα ενδεχόμενα;

# ΕΝΟΤΗΤΑ 1.2. ΠΙΘΑΝΟΤΗΤΕΣ: ΟΡΙΣΜΟΙ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

## Διερεύνηση

***Πόσο πιθανό είναι να συμβει;***

Παρακάτω περιγράφονται δύο παιχνίδια για δύο παίκτες.

**Παιχνίδι 1:** Οι δύο παίκτες ο Α και ο Β στρίβουν ένα κέρμα δύο φορές ο καθένας και καταγράφουν τα αποτελέσματα. Ο παίκτης Α κερδίζει αν έρθει δύο φορές γράμματα. Ο παίκτης Β κερδίζει αν έρθει διαφορετικό αποτέλεσμα σε κάθε ρίψη.

**Παιχνίδι 2:** Στο διπλανό σχήμα φαίνεται ένας δίσκος χωρισμένος σε τρεις ίσους κυκλικούς τομείς, από τους οποίους καθένας έχει άλλο χρώμα και έναν διαφορετικό αριθμό επάνω του. Το βέλος περιστρέφεται και σταματά. Αν σταμάτησε σε θετικό αριθμό κερδίζει ο παίκτης Α, ενώ αν σταμάτησε σε αρνητικό αριθμό κερδίζει ο παίκτης Β.


— Θεωρείτε ότι τα παραπάνω παιχνίδια είναι δίκαια και για τους δύο παίκτες;
— Πώς το εξηγείτε;
— Αν όχι, τι θα μπορούσε να αλλάξει ώστε να γίνουν;

## Βασικές μαθηματικές έννοιες -Ιδέες-Διεργασίες

***Εισαγωγικά για την έννοια της πιθανότητας***

Τα παιχνίδια που περιγράφονται παραπάνω μπορούν να θεωρηθούν ως πειράματα τύχης, για τα οποία μιλήσαμε στην προηγούμενη παράγραφο, καθώς ένα κύριο χαρακτηριστικό τους είναι πως δεν μπορεί να προβλεφθεί το αποτέλεσμά τους. Δηλαδή υπάρχει αβεβαιότητα σχετικά με το ποιος παίκτης θα κερδίσει σε κάθε παιχνίδι. Η αβεβαιότητα αυτή όμως είναι ίδια και στις δύο περιπτώσεις παιχνιδιών; Δηλαδή, αν θεωρήσουμε το ενδεχόμενο «κερδίζει ο παίκτης Α», η βεβαιότητα για την πραγματοποίηση του είναι η ίδια και στα δύο παιχνίδια; Πώς μπορούμε να «μετρήσουμε» τη βεβαιότητα αυτή;

Η πιθανότητα ενός ενδεχομένου, όπως το προηγούμενο, είναι ένας αριθμός που εκφράζει το μέτρο της βεβαιότητας που αποδίδουμε στο να πραγματοποιηθεί το ενδεχόμενο αυτό.

Στην ειδική περίπτωση που θεωρούμε ότι τα δυνατά αποτελέσματα ενός πειράματος τύχης είναι ισοπίθανα, όπως στην ρίψη ενός αμερόληπτου κέρματος ή ζαριού, ισχύει ο παρακάτω ορισμός.

**Κλασικός ορισμός πιθανότητας**

Σε ένα πείραμα τύχης με ν ισοπίθανα αποτελέσματα, η πιθανότητα ενός ενδεχομένου Α που περιέχει κ τέτοια αποτελέσματα είναι:

$$P(A)=\frac{\text{Πλήθος ευνοϊκών αποτελεσμάτων για το Α}}{\text{Πλήθος όλων των δυνατών αποτελεσμάτων}}=\frac{\kappa}{\nu}$$

Άμεσες συνέπειες του κλασικού ορισμού είναι οι παρακάτω:

— $P(\Omega) = 1$,

— $P(\varnothing) = 0$,

— Για κάθε ενδεχόμενο του δ.χ. Ω ισχύει $0 \le P(A) \le 1$

Για παράδειγμα, ας θεωρήσουμε το παράδειγμα της ρίψης ενός αμερόληπτου κέρματος με δ.χ.:

$$\Omega =\{1, 2, 3, 4, 5, 6\}$$

και το ενδεχόμενο Α: το αποτέλεσμα της ρίψης είναι μεγαλύτερο του 4, δηλαδή:

$$A =\{5, 6\}$$

Τότε η πιθανότητα να πραγματοποιηθεί το Α είναι:

$$P(A)=\frac{2}{6}$$

Στην προηγούμενη παράγραφο, θεωρώντας έναν δ.χ. ενός πειράματος τύχης, αυτό που κάναμε είναι ένα μέρος της μοντελοποίησης του πειράματος, λέγοντας τι μπορεί να συμβεί. Τώρα εμπλουτίζουμε το μοντέλο μας, αφού μπορούμε να πούμε και πόσο πιθανά θεωρούμε να συμβούν αυτά που μπορούν να συμβούν.

## Εφαρμογές

### Εφαρμογή 1

Ρίχνουμε δύο αμερόληπτα ζάρια και καταγράφουμε το άθροισμα των ενδείξεών τους. Ποιο άθροισμα έχει μεγαλύτερη πιθανότητα εμφάνισης;

**Λύση**

Μπορούμε να καταγράψουμε τα πιθανά αποτελέσματα όπως φαίνεται στον πίνακα. Στη συνέχεια θεωρούμε τα 36 διατεταγμένα ζεύγη στα κελιά του πίνακα ως στοιχεία του δ.χ. της ρίψης δύο ζαριών. Για δύο πραγματικά αμερόληπτα ζάρια είναι ρεαλιστικό να υποθέσουμε ότι τα 36 αυτά αποτελέσματα είναι εξίσου πιθανά.

*Ο δειγματικός χώρος της ρίψης δύο ζαριών που χρησιμοποιούμε στην εφαρμογή 1*

| | | | | | |
|---|---|---|---|---|---|
| 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 |

Σύμφωνα με τον κλασικό ορισμό της πιθανότητας, το ενδεχόμενο $A_7$: το άθροισμα των δύο ρίψεων είναι ίσο με 7, δηλαδή το

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

έχει πιθανότητα να πραγματοποιηθεί:

$$P(A_7) = \frac{6}{36} = \frac{1}{6}$$

Αυτό είναι και το άθροισμα με τη μεγαλύτερη πιθανότητα εμφάνισης. Οποιοδήποτε άλλο άθροισμα έχει μικρότερη πιθανότητα εμφάνισης, καθώς τα ευνοϊκά ενδεχόμενα είναι λιγότερα από 6.

Π.χ. το ενδεχόμενο $A_6$: το άθροισμα των δύο ρίψεων είναι ίσο με 6

$$A_6 = \{(5,1), (4,2), (3,3), (2,4), (2,5), (1,5)\}$$

έχει πιθανότητα ίση με:

$$P(A_6) = \frac{5}{36} < \frac{1}{6}$$

**Χρήση του κλασικού ορισμού**

## Διερεύνηση

**Παιχνίδι 3:** Στο διπλανό σχήμα φαίνεται ένας δίσκος χωρισμένος σε έναν κίτρινο και έναν μπλε τομέα. Ο παίκτης Α περιστρέφει μία φορά το βέλος. Κερδίζει αν το βέλος βρεθεί σε μπλε περιοχή, αλλιώς κερδίζει ο Β. Ποια είναι η πιθανότητα να κερδίσει ο παίκτης Α;

## Βασικές μαθηματικές έννοιες -Ιδέες-Διεργασίες

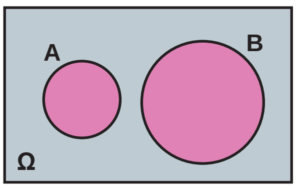
Είναι ιδιαίτερα συνηθισμένο τα δυνατά αποτελέσματα ενός πειράματος τύχης να μην είναι ισοπίθανα. Υπάρχουν απλά παραδείγματα όπως το γνωστό παιχνίδι όπου κάποιος πετάει στον αέρα ένα μπουκάλι μισογεμάτο με νερό και καταγράφει αν θα πέσει όρθιο ή όχι στο έδαφος (bottle flip), αλλά και πιο σύνθετα όπως οι λειτουργίες κυκλωμάτων υπολογιστών και κινητών τηλεφώνων. Σε αυτή την κατηγορία εμπίπτουν τα περισσότερα από τα ενδιαφέροντα προβλήματα που μελετά η Θεωρία των Πιθανοτήτων. Σε τέτοιες περιπτώσεις ο κλασικός ορισμός της πιθανότητας δεν μπορεί να χρησιμοποιηθεί, μπορεί όμως να χρησιμοποιηθεί ο αξιωματικός ορισμός, που εφαρμόζεται ευρύτερα από τον κλασικό.

Στόχος μας είναι να κατασκευάσουμε μία αντιστοιχία μεταξύ των ενδεχομένων ενός δ.χ. ενός πειράματος τύχης και πραγματικών αριθμών που να ανήκουν στο διάστημα [0, 1], έτσι ώστε κάθε αριθμός να εκφράζει πόσο πιθανό θεωρούμε να πραγματοποιηθεί το αντίστοιχο ενδεχόμενο. Παρακάτω περιγράφεται πώς μπορεί να κατασκευαστεί αυτή η αντιστοιχία.

***Η αντιστοιχία που ορίζεται με τον αξιωματικό ορισμό***

***Α→P(A)***

***Ο απλός προσθετικός νόμος***




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

### Αξιωματικός ορισμός πιθανότητας

Έστω $\Omega=\{\omega_1, \omega_2,...,\omega_\nu\}$ ένας δ.χ. ενός πειράματος τύχης. Σε κάθε ενδεχόμενο Α του Ω **αποδίδουμε** έναν πραγματικό αριθμό που ονομάζουμε πιθανότητα του Α και συμβολίζουμε με P(A), έτσι ώστε:

— $P(A) \geq 0$, για οποιοδήποτε Α του Ω

— $P(\Omega) = 1$

— Ικανοποιείται ο απλός προσθετικός νόμος:

Αν τα ενδεχόμενα Α και Β είναι οποιαδήποτε ασυμβίβαστα ενδεχόμενα του Ω, τότε $P(A \cup B) = P(A) + P(B)$.

Θα δούμε στην επόμενη παράγραφο ότι ο αξιωματικός ορισμός έχει τις παρακάτω συνέπειες:

$$P(A) \leq 1, \text{ για οποιοδήποτε ενδεχόμενο Α και } P(\varnothing) = 0.$$

Η πιθανότητα P(A), δηλαδή ο αριθμός που αποδίδουμε στο Α και που ικανοποιεί τα τρία αξιώματα του ορισμού, είναι ένας αριθμός ανάμεσα στο 0 και στο 1 και εκφράζει το πόσο πιθανό θεωρούμε να συμβεί το Α. Στο αδύνατο ενδεχόμενο αποδίδουμε πιθανότητα 0 και στο βέβαιο ενδεχόμενο (το Ω) αποδίδουμε πιθανότητα 1.

Για να μελετήσουμε, με χρήση του αξιωματικού ορισμού, ένα πείραμα τύχης κάνουμε το εξής:

Προσδιορίζουμε τα δυνατά αποτελέσματα $\omega_1, \omega_2,\ldots,\omega_\nu$ δηλαδή έναν κατάλληλο δ.χ. για το πείραμα τύχης και αντιστοιχίζουμε στα απλά ενδεχόμενα $\{\omega_1\}, \{\omega_2\},\ldots, \{\omega_\nu\}$ τους πραγματικούς αριθμούς $p_1, p_2,\ldots,p_\nu$ που εκφράζουν πόσο πιθανό θεωρούμε κάθε δυνατό αποτέλεσμα. Δηλαδή $P(\{\omega_i\})=p_i$ για κάθε $i = 1,2,3,\ldots,\nu$.

Αυτό το κάνουμε ώστε να ισχύουν:

— $0 \leq p_i$, για κάθε $i = 1,2,\ldots,\nu$ (αφού θέλουμε $P(A)\geq 0$ )

— $p_1 + p_2 +\ldots+p_\nu = 1$ (αφού θέλουμε $P(\Omega)=1$)

Στη συνέχεια, προκειμένου να αποδώσουμε πιθανότητα σε οποιοδήποτε άλλο ενδεχόμενο χρησιμοποιούμε το 3ο αξίωμα (τον απλό προσθετικό νόμο). Π.χ.

$$P(\{\omega_1, \omega_2, \omega_3\}) = P(\{\omega_1\}) + P(\{\omega_2\}) + P(\{\omega_3\}) = p_1 + p_2 + p_3$$

Με αυτόν τον τρόπο αποδίδουμε πιθανότητα σε οποιοδήποτε σύνθετο ενδεχόμενο του Ω, εξασφαλίζοντας ότι ο αξιωματικός ορισμός ικανοποιείται.

Ο κλασικός ορισμός της πιθανότητας είναι υποπερίπτωση του αξιωματικού ορισμού. Στην εφαρμογή 1, για παράδειγμα, θεωρούμε ως απλά ενδεχόμενα του δ.χ. τα $\{\omega_i\}$, για i=1, 2, ..., 36, όπου κάθε $\omega_i$ αντιστοιχεί σε ένα κελί του πίνακα που εκφράζει τον δ.χ. της ρίψης των δύο ζαριών.

Θεωρώντας -ρεαλιστικά- τα αποτελέσματα $\omega_i$ ισοπίθανα, τότε σε κάθε $\{\omega_i\}$ με i = 1,2,…,36 αποδίδουμε πιθανότητα $p_i = \frac{1}{36} \geq 0$.

Επίσης ισχύει $p_1 + p_2 + ... + p_{36} = 36 \cdot \frac{1}{36} = 1$.

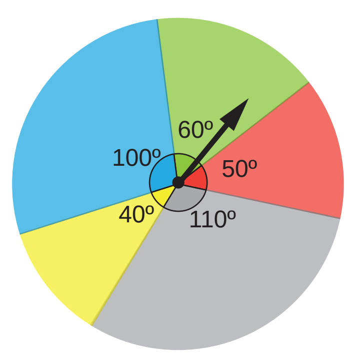
Όπως είδαμε, εφόσον ισχύουν τα παραπάνω, μπορούμε να αποδώσουμε πιθανότητα σε οποιοδήποτε σύνθετο ενδεχόμενο του Ω, εξασφαλίζοντας ότι ικανοποιείται ο αξιωματικός ορισμός.

Έτσι, αν π.χ. $A = \{\omega_1, \omega_2, \dots, \omega_\kappa\}$ τότε $P(A) = p_1 + p_2 + ... + p_\kappa = \frac{\kappa}{36}$, που είναι ακριβώς το ίδιο με αυτό που λέει ο κλασικός ορισμός. Άρα οι δύο ορισμοί είναι ισοδύναμοι στην περίπτωση αυτή.

## Εφαρμογές

### Εφαρμογή 2

Στο παρακάτω σχήμα φαίνεται ένας δίσκος, χωρισμένος σε χρωματισμένους κυκλικούς τομείς. Ο δίσκος χρησιμεύει σε ένα παιχνίδι με τους εξής κανόνες. Κάθε παίκτης επιλέγει ένα χρώμα και ένας από όλους περιστρέφει το βέλος με δύναμη. Το βέλος σταματάει την περιστροφή του. Νικητής είναι ο παίκτης που είχε επιλέξει το χρώμα του τομέα που σταμάτησε το βέλος.




Αν το βέλος σταματήσει μεταξύ δύο τομέων, τότε θεωρούμε ότι βρίσκεται στον τομέα που είναι το μεγαλύτερο μέρος του. Αν σταματήσει έτσι ώστε να μην είναι φανερό σε ποιον τομέα βρίσκεται το μεγαλύτερο μέρος του, η περιστροφή του βέλους επαναλαμβάνεται.

**α)** Η Γιάννα και ο Λεωνίδας προσπάθησαν να μοντελοποιήσουν το πείραμα τύχης ορίζοντας δ.χ. $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ με τα απλά ενδεχόμενα:

$\omega_1$: «το βέλος σταματάει στο πράσινο»

$\omega_2$: «το βέλος σταματάει στο γαλάζιο»

$\omega_3$: «το βέλος σταματάει στο κίτρινο»

$\omega_4$: «το βέλος σταματάει στο γκρι»

$\omega_5$: «το βέλος σταματάει στο κόκκινο»

Όμως ακολούθησαν δύο διαφορετικούς τρόπους:

Η Γιάννα αντιστοίχισε πιθανότητα $p_i = \frac{1}{5}$, σε κάθε ένα από τα $\{\omega_i\}$.

Ο Λεωνίδας αντιστοίχισε πιθανότητα $p_i$ ίση με τις μοίρες του τομέα του αντίστοιχου χρώματος. Π.χ. $p_1 = 60$.
Να σχολιάσετε τους δύο τρόπους μοντελοποίησης.

**β)** Να βρείτε τις πιθανότητες να κερδίσει κάθε παίκτης, σε σχέση με το χρώμα που επέλεξε.

**<u>Λύση</u>**

**α) Γιάννα:** Σε αυτή την περίπτωση η αντιστοιχία που ορίζεται είναι συνεπής με τα τρία αξιώματα του αξιωματικού ορισμού των πιθανοτήτων.
Όμως, από την άλλη μεριά, αυτή η αντιστοίχιση δεν εκφράζει την αντίληψή μας για τις πιθανότητες έκβασης του συγκεκριμένου πειράματος τύχης. Π.χ.

$$p_1 = p_2 = \frac{1}{5}$$

Ενώ αυτό που περιμένουμε να ισχύει είναι:

$$p_2 > p_1$$

καθώς ο γαλάζιος τομέας αντιστοιχεί σε μεγαλύτερη επίκεντρη γωνία.
Συνεπώς, δε θα μπορούσε να μοντελοποιηθεί ρεαλιστικά το πείραμα τύχης με τον τρόπο αυτό.
**Λεωνίδας:** Σε αυτή την περίπτωση η αντιστοιχία που ορίζεται μπορεί να είναι συμβατή με την αντίληψή μας για την έκβαση του πειράματος τύχης, ωστόσο η μοντελοποίηση είναι προβληματική, γιατί:

— Αν μετρούσαμε τη γωνία σε rad, η τιμή της πιθανότητας κάθε απλού ενδεχομένου θα άλλαζε, ενώ η πιθανότητα πρέπει να είναι ανεξάρτητη από μονάδες μέτρησης.

— Δεν είναι συμβατή με την απαίτηση $P(\Omega)=1$ του αξιωματικού ορισμού καθώς προκύπτει $P(\Omega)=360$.

Στην πράξη, ορίζοντας με αυτό τον τρόπο την αντιστοιχία δεν μπορεί να απαντηθεί το εξής ερώτημα:

«Τι είναι πιθανότερο από τα επόμενα; Να σταματήσει το βέλος στο πράσινο χρώμα ή να έρθει άθροισμα 7, μετά από ρίψη δύο ζαριών;»
Συνεπώς, και αυτός ο τρόπος μοντελοποίησης είναι λανθασμένος.

**Χρήση του αξιωματικού ορισμού**

**β)** Για να απαντήσουμε πρέπει να μοντελοποιήσουμε σωστά το πείραμα τύχης. Αν θεωρήσουμε ότι το βέλος περιστρέφεται χωρίς να «φρενάρει» ξαφνικά πάνω από κάποιο χρώμα, δηλαδή η περιστροφή του γίνεται με «φυσιολογικό» τρόπο (ή αλλιώς αμερόληπτα), τότε αναμένουμε η πιθανότητα του ενδεχομένου «Το βέλος σταματάει στο χρώμα Χ» να είναι ανάλογη της επίκεντρης γωνίας του τομέα χρώματος Χ.
Έτσι μοντελοποιούμε το πείραμα τύχης με έναν δ.χ.

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$$

ώστε κάθε ένα από τα απλά ενδεχόμενα $\{\omega_i\}$ να αντιστοιχεί σε ένα από τα πέντε χρώματα του δίσκου, όπως στην εκφώνηση, και ορίζουμε την πιθανότητα $p_i$ ως εξής:

$$p_i = \frac{\text{οι μοίρες του τομέα που αντιστοιχεί στο } \omega_i}{360^o}$$

Π.χ. Η πιθανότητα του ενδεχομένου «το βέλος σταματάει στον πράσινο τομέα» είναι:

$$p_1 = \frac{60^o}{360^o} = \frac{1}{6}$$

Στη συνέχεια η πιθανότητα οποιουδήποτε ενδεχομένου του πειράματος τύχης μπορεί να υπολογιστεί με βάση τα $p_i$.
Π.χ. αν Α: «το βέλος σταματάει στο πράσινο ή στο γαλάζιο», τότε:

$$A = \{\omega_1\} \cup \{\omega_2\}$$

Άρα:

$$P(A) = p_1 + p_2 = \frac{1}{6} + \frac{5}{18} = \frac{8}{18} = \frac{4}{9}$$

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Ρίχνουμε δύο αμερόληπτα ζάρια και καταγράφουμε το αποτέλεσμα των ρίψεων. Ποιες είναι οι πιθανότητες των ενδεχομένων:
**α)** Το άθροισμα των ρίψεων είναι ίσο με 4.
**β)** Το άθροισμα των ρίψεων είναι μεγαλύτερο του 4.
**γ)** Το άθροισμα των ρίψεων είναι περιττός αριθμός.

**2)** Να αποδείξετε ότι ο απλός προσθετικός νόμος προκύπτει ως συνέπεια του κλασικού ορισμού.

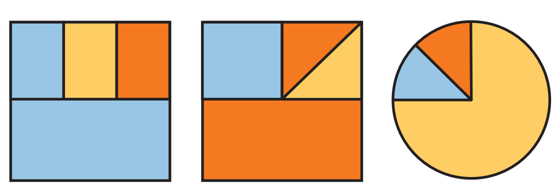
**3)** Ένα κέρμα είναι κατασκευασμένο έτσι ώστε κατά τη ρίψη του η πιθανότητα του ενδεχομένου «εμφανίζεται κεφαλή» είναι 0,95. Θεωρείτε ότι το πείραμα αυτό είναι πείραμα τύχης;

**4)** Ρίχνουμε δύο αμερόληπτα ζάρια και καταγράφουμε το αποτέλεσμα των ρίψεων. Να αντιστοιχίσετε τα ενδεχόμενα της 1ης στήλης με τις πιθανότητες της 2ης στήλης:

| 1η στήλη | 2η στήλη |
|---|---|
| — Έρχεται διπλή ζαριά (το ίδιο αποτέλεσμα και στα δύο ζάρια.) | $\frac{1}{6}$ |
| — Το αποτέλεσμα του ενός μόνο ζαριού είναι περιττός αριθμός. | $\frac{1}{4}$ |
| — Το αποτέλεσμα του ενός μόνο ζαριού είναι άρτιος αριθμός. | $\frac{3}{4}$ |
| — Τουλάχιστον ένα από τα δύο ζάρια φέρνει άρτιο αποτέλεσμα. | 0,5 |
| — Το αποτέλεσμα και των δύο ζαριών είναι άρτιος αριθμός. | |
| — Το άθροισμα των αποτελεσμάτων των δύο ζαριών είναι άρτιος αριθμός. | |

**5)** Να μοντελοποιήσετε το πείραμα τύχης της εφαρμογής 2, χρησιμοποιώντας το μέτρο του κάθε τόξου σε rad αντί σε μοίρες, για να ορίσετε τις πιθανότητες των στοιχειωδών ενδεχομένων $\omega_i$. Να συγκρίνετε τους δύο τρόπους μοντελοποίησης. Τι παρατηρείτε;

**6)** Κάθε ένα από τα παρακάτω τρία σχήματα (δύο τετράγωνα και ένα κυκλικό) εμφανίζεται στην οθόνη ενός από τρεις ίδιους ηλεκτρονικούς υπολογιστές.



Πλαίσιο 1 Πλαίσιο 2 Πλαίσιο 3

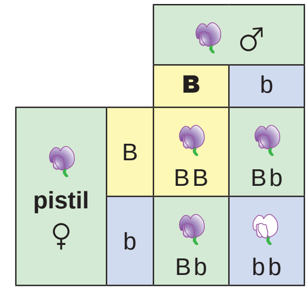
Ένα πρόγραμμα στον υπολογιστή επιλέγει τυχαία και χρωματίζει με μαύρο χρώμα ένα pixel μέσα στο σχήμα. Σε κάθε πλαίσιο ποια είναι η πιθανότητα του ενδεχομένου το pixel που θα χρωματιστεί:

α) Να ήταν κόκκινο;

β) Να ήταν γαλάζιο ή κίτρινο;

γ) Να μην ήταν κόκκινο;

**7)** Ο βοτανολόγος Γκρέγκορ Μέντελ, στα μέσα του 19ου αιώνα, πειραματίστηκε με την εμφάνιση των κληρονομικών χαρακτηριστικών των μοσχομπίζελων, διατυπώνοντας τους νόμους της Μενδελικής κληρονομικότητας. Από τα πειράματά του συμπέρανε τον νόμο της ομοιομορφίας, σύμφωνα με τον οποίο το χρώμα του άνθους μοσχομπίζελου είναι αποτέλεσμα του συνδυασμού δύο «κληρονομικών παραγόντων» που σήμερα ονομάζονται αλληλόμορφα γονίδια. Για το χρώμα υπάρχουν δύο γονίδια: το επικρατές B, που αντιστοιχεί στο ιώδες και το υπολειπόμενο b που αντιστοιχεί στο λευκό χρώμα.




Σε ένα φυτό, το χρώμα του άνθους του οφείλεται σε ένα αλληλόμορφο γονίδιο που κληρονομεί από το πατρικό μοσχομπίζελο κι ένα αλληλόμορφο γονίδιο που κληρονομεί από το μητρικό. Όπως φαίνεται στον πίνακα για να προκύψει μοσχομπίζελο με λευκό άνθος, πρέπει και τα δύο αλληλόμορφα γονίδια να είναι τύπου b. Σε κάθε άλλη περίπτωση προκύπτει μοσχομπίζελο με ιώδες άνθος.

Διασταυρώνουμε δύο μοσχομπίζελα: ένα τύπου BB και ένα με λευκό άνθος.
Να υπολογίσετε την πιθανότητα να προκύψει μοσχομπίζελο με λευκό άνθος:
α) στην 1η (θυγατρική) γενιά.
β) στην 2η (θυγατρική) γενιά.

**8)** Σε ένα μαιευτήριο τον προηγούμενο μήνα από τα παιδιά που γεννήθηκαν το 30% ήταν αγόρια και το 70% ήταν κορίτσια. Με ποιες από τις παρακάτω προτάσεις συμφωνείτε;

**α)** Σε αυτό το μαιευτήριο τον προηγούμενο μήνα γεννήθηκαν περισσότερα κορίτσια.

**β)** Αν ένα ζευγάρι που περιμένει παιδί επιλέξει αυτό το μαιευτήριο για τον τοκετό, τότε η πιθανότητα να γεννηθεί κορίτσι είναι μεγαλύτερη από την πιθανότητα να γεννηθεί αγόρι.

**γ)** Αν επιλέξουμε τυχαία ένα παιδί από τον κατάλογο των νεογέννητων του προηγούμενου μήνα σε αυτό το μαιευτήριο, τότε η πιθανότητα να είναι αγόρι είναι ίση με 0,3.

**δ)** Αν επιλέξουμε ένα παιδί στην τύχη από την λίστα των παιδιών που έχουν γεννηθεί σε αυτό το μαιευτήριο, τότε η πιθανότητα να είναι κορίτσι είναι 0,7.

**9)** Σε μία κλειστή κάλπη τοποθετούνται 5 κόκκινα και 6 πράσινα σφαιρίδια. Από την κάλπη βγάζουμε τυχαία ένα σφαιρίδιο. Αφού βγάλουμε το σφαιρίδιο, ποια είναι η πιθανότητα του ενδεχομένου να υπάρχει στην κάλπη ίδιος αριθμός σφαιριδίων από τα δύο χρώματα; Αν στην κάλπη αρχικά υπήρχαν περισσότερα σφαιρίδια, αλλά πάλι τα πράσινα ήταν περισσότερα από τα κόκκινα κατά 1, να διερευνήσετε αν και πόσο θα άλλαζε η πιθανότητα του ίδιου ενδεχομένου;

**10)** Ρίχνουμε ένα αμερόληπτο κέρμα δύο φορές και καταγράφουμε το αποτέλεσμα. Ορίζουμε τα παρακάτω ενδεχόμενα:
**Α:** έρχεται το πολύ μία φορά Κεφαλή.
**Β:** έρχεται τουλάχιστον μία φορά Κεφαλή.
**Γ:** το αποτέλεσμα των δύο ρίψεων είναι διαφορετικό.

**Δ:** το αποτέλεσμα των δύο ρίψεων είναι το ίδιο.

**α)** Να αποδείξετε ότι $P(A)=P(B)$ και ότι $P(\Gamma) = P(\Delta)$.

**β)** Να υπολογίσετε τις πιθανότητες των ενδεχομένων $A \cup B$, $A \cap B$, $A - B$.

**γ)** Να υπολογίσετε τις πιθανότητες των ενδεχομένων $\Gamma'$, $\Gamma \cap \Delta$, $\Gamma \cup \Delta$, $B \cup \Gamma'$.

## Πρόσθετο Υλικό

### Ερωτήματα για διερεύνηση

**1)** Από τον ορισμό της πιθανότητας (κλασικό και αξιωματικό) γνωρίζετε ότι $A = \varnothing \Rightarrow P(A) = 0$. Ισχύει το αντίστροφο; Μπορεί να υπάρχει ενδεχόμενο Α ενός δ.χ., που να μην είναι κενό και να ισχύει $P(A) = 0$ ;

**2)** **α)** Παρακάτω περιγράφεται ένα παιχνίδι (πείραμα τύχης) με αμερόληπτο κέρμα, για δύο παίκτες, την Άννα και τον Βασίλη.
Η Άννα κάνει 2 ρίψεις του κέρματος, στη συνέχεια κάνει 1 ρίψη ο Βασίλης και καταγράφεται το αποτέλεσμα των ρίψεων. Η Άννα κερδίζει αν φέρνει περισσότερες κεφαλές (Κ) από τον Βασίλη. Σε όλες τις άλλες περιπτώσεις κερδίζει ο Βασίλης. Ποια είναι η πιθανότητα να κερδίσει η Άννα; Είναι δίκαιο το παιχνίδι;
**β)** Τι θεωρείτε ότι θα συνέβαινε στην πιθανότητα να κερδίσει η Άννα στο προηγούμενο παιχνίδι, αν γίνονταν περισσότερες ρίψεις του νομίσματος (η Άννα κάνει ν+1 ρίψεις και ο Βασίλης ν ρίψεις);

**3)** Αφού διαβάσετε το παρακάτω απόσπασμα να διερευνήσετε τα ερωτήματα (α) και (β), που ακολουθούν.

***Συνθετική εργασία***

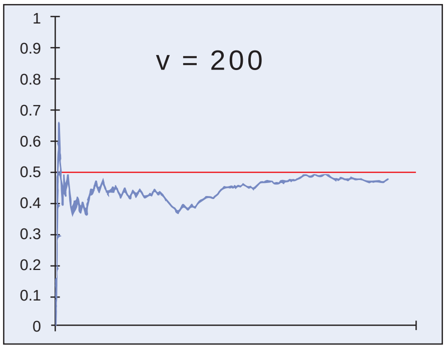
*Ο νόμος των μεγάλων αριθμών*

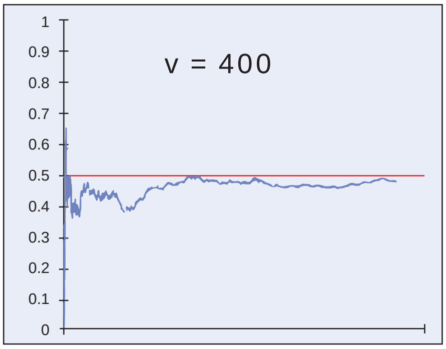
*Σε έναν πίνακα καταγράφονται τα αποτελέσματα της εκτέλεσης του ακόλουθου πειράματος:*

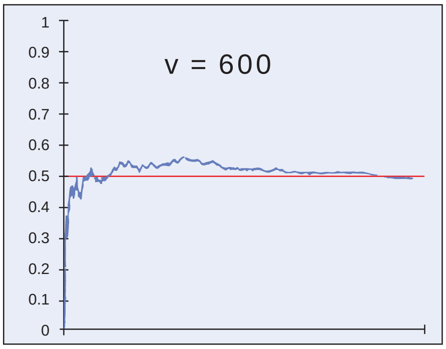
| ν | κ | $\frac{\kappa}{\nu}$ |
|---|---|---|
| 10 | 7 | 0,7 |
| 20 | 13 | 0,65 |
| 30 | 16 | 0,533 |
| 40 | 23 | 0,575 |
| 50 | 26 | 0,52 |
| … | … | … |
| 180 | 89 | 0,494 |
| 190 | 93 | 0,489 |
| 200 | 99 | 0,495 |

*Ρίχνουμε ένα συμμετρικό και ομογενές νόμισμα και σημειώνουμε με Κ το αποτέλεσμα “κεφαλή” και με Γ το αποτέλεσμα “γράμματα”. Στο διπλανό απόσπασμα του πίνακα φαίνονται το πλήθος των ρίψεων (ν), το πλήθος εμφάνισης Κ (κ) και ο λόγος $\frac{\kappa}{\nu}$. Παρατηρούμε ότι καθώς αυξάνεται ο αριθμός ν των ρίψεων ο λόγος $\frac{\kappa}{\nu}$ σταθεροποιείται γύρω από την τιμή 0,5 ή, όπως λέμε, “τείνει” στον αριθμό 0,5.*

*Αντίστοιχα, το ίδιο φαίνεται και στα παρακάτω διαγράμματα.*




*Στον οριζόντιο άξονα παριστάνεται το πλήθος των ρίψεων (ν) και στον κατακόρυφο άξονα η τιμή του λόγου* $\frac{\kappa}{\nu}$*, όπου κ είναι το πλήθος εμφάνισης Κ (όπως αναφέρεται και παραπάνω). Το διάγραμμα πάνω αριστερά αντιστοιχεί σε μία αλληλουχία 200 ρίψεων, το πάνω δεξιά σε αλληλουχία 400 ρίψεων και το κάτω διάγραμμα σε αλληλουχία 600 ρίψεων.*

*Αυτό επιβεβαιώνει την «αντίληψή» μας για το αποτέλεσμα της ρίψης ενός «αμερόληπτου» νομίσματος. Σε ανάλογα συμπεράσματα οδηγούμαστε και με άλλα παραδείγματα, όπως η ρίψη του ζαριού, που εκεί η συχνότητα εμφάνισης κάθε όψης τείνει στο* $\frac{1}{6}$*. Το εμπειρικό αυτό εξαγόμενο, το οποίο αποδεικνύεται και θεωρητικά, ονομάζεται στατιστική ομαλότητα ή νόμος των μεγάλων αριθμών.*

**α)** Στο διπλανό σχήμα φαίνεται ένας δίσκος που λειτουργεί όπως ο δίσκος της εφαρμογής 2 και χρησιμοποιείται για ένα αντίστοιχο παιχνίδι. Τρία παιδιά, η Άννα, η Ντενίζ και ο Κυριάκος πριν παίξουν, θέλουν να έχουν μία «εικόνα» για το πόσο πιθανό είναι το βέλος να σταματήσει σε κάθε χρώμα. Ωστόσο, δε γνωρίζουν αν το βέλος περιστρέφεται αμερόληπτα. Αποφάσισαν, λοιπόν, να κάνουν 100 δοκιμές περιστρέφοντας το βέλος και να καταγράψουν πόσες φορές σταματά πάνω από κάθε χρώμα.

Βρήκαν:

| Χρώμα που σταμάτησε το βέλος | Πόσες φορές |
|---|---|
| Κίτρινο | 18 |
| Μαύρο | 20 |
| Μπλέ | 17 |
| Πράσινο | 17 |
| Κόκκινο | 15 |
| Καφέ | 13 |

Η Άννα ισχυρίζεται ότι η πιθανότητα να σταματήσει το βέλος στο κίτρινο είναι 0,18. Ο Κυριάκος θεωρεί ότι το βέλος δεν περιστρέφεται αμερόληπτα, γιατί το εμβαδόν του κίτρινου τομέα είναι πολύ μεγαλύτερο του εμβαδού του μαύρου τομέα. Η Ντενίζ προτείνει ότι δεν είναι ασφαλές να βγάλουν συμπέρασμα με 100 στριψίματα και ότι χρειάζεται να στρίψουν το βέλος πολλές φορές ακόμα. Πώς σχολιάζετε τις ενέργειες και τις εικασίες των παιδιών;

**β)** Μία παρέα παιδιών η Ελένη, ο Χασάν, ο Βαγγέλης και η Γιάννα θέλουν να παίξουν με ένα μισογεμάτο μπουκάλι. Το πετάει κάποιος από όλους στον αέρα, ενώ προηγουμένως έχουν όλοι δηλώσει αν θα σταθεί όρθιο ή όχι. Όποιος «μαντέψει» σωστά, κερδίζει. Για να περιγράψουν πόσο πιθανό είναι να σταθεί το μπουκάλι όρθιο, τα παιδιά σκέφτηκαν το εξής: Θα πετάξουν 100 φορές το μπουκάλι στον αέρα και θα καταγράψουν πόσες φορές θα σταθεί όρθιο. Αν σταθεί ν φορές όρθιο τότε θα ορίσουν την πιθανότητα του ενδεχομένου «το μπουκάλι στέκεται όρθιο» ίση με $\frac{\nu}{100}$ και του ενδεχομένου «το μπουκάλι δε στέκεται όρθιο» ίση με $1-\frac{\nu}{100}$.

Πώς σχολιάζετε την εικασία των παιδιών;

***Συνθετική εργασία***

*Ο Γκρέγκορ Γιόχαν Μέντελ ήταν Αυστριακός μοναχός και βοτανολόγος. Έκανε τα πειράματά του στο μοναστήρι που ζούσε, στο Μπρυν της Αυστροουγγαρίας. Η δημοσίευση των συμπερασμάτων του δεν έτυχε προσοχής από την επιστημονική κοινότητα της εποχής με αποτέλεσμα ο Μέντελ να επηρεαστεί και να εγκαταλείψει τα πειράματά του για πάντα, ασχολούμενος με τη διοίκηση του μοναστηριού όπου ζούσε.*

**4)** Ανάμεσα στα έτη 1856 και 1863, ο Μέντελ καλλιέργησε και μελέτησε περίπου 28.000 μπιζελιές διατυπώνοντας τους νόμους του για την κληρονομικότητα.

**α)** Να ερευνήσετε γιατί ο Μέντελ χρησιμοποίησε μοσχομπίζελα στα πειράματά του.

**β)** Ένας σύγχρονος βοτανολόγος επανέλαβε το πείραμα του Μέντελ ξεκινώντας (πατρική γενιά) από μοσχομπίζελα τύπου ΒΒ και μοσχομπίζελα με λευκό άνθος, ώστε όσα είναι τα μεν να είναι και τα δε. Τελικά, όταν έφτασε στην 2η θυγατρική γενιά μοσχομπίζελων υπήρχαν 433 ιώδη και 170 λευκά μοσχομπίζελα. Να σχολιάσετε τα αποτελέσματα του πειράματος.

Προτεινόμενες ασκήσεις από την παράγραφο 1.2 του βιβλίου «Άλγεβρα και Στοιχεία Πιθανοτήτων» της Α΄ Λυκείου:
Α΄ ομάδας: 1, 2, 3, 4, 5, 6.
Ερωτήσεις κατανόησης 1ου κεφαλαίου: 1, 2.

# ΕΝΟΤΗΤΑ 1.3. ΠΙΘΑΝΟΤΗΤΕΣ ΚΑΙ ΠΡΑΞΕΙΣ ΜΕ ΕΝΔΕΧΟΜΕΝΑ

## Διερεύνηση

Οι μαθητές/τριες ενός Λυκείου έχουν τη δυνατότητα να συμμετέχουν στις παρακάτω δραστηριότητες:

— Θεατρική ομάδα
— Ομάδα στίβου
— Όμιλος μουσικής

Επιλέγουμε τυχαία έναν από τους μαθητές/τριες του Λυκείου. Έστω ότι η πιθανότητα να συμμετέχει ο/η μαθητής/τρια που επιλέγουμε:

— στη θεατρική ομάδα είναι ίση με $\frac{2}{5}$,
— στην ομάδα στίβου είναι ίση με $\frac{7}{15}$,
— στον όμιλο μουσικής είναι ίση με $\frac{1}{4}$.

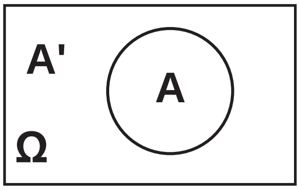
**α)** Θεωρείτε ότι σε αυτό το σχολείο απαγορεύεται ένας/μία μαθητής/τρια να συμμετέχει σε πάνω από μία δραστηριότητες;
**β)** Αν η πιθανότητα του ενδεχομένου «επιλέγουμε έναν/μία μαθητή/τρια που συμμετέχει στη θεατρική ομάδα ή στην ομάδα στίβου» είναι μικρότερη από $\frac{13}{15}$, τι μπορούμε να συμπεράνουμε;
**γ)** Αν η πιθανότητα του ενδεχομένου «επιλέγουμε έναν/μία μαθητή/τρια που συμμετέχει στη θεατρική ομάδα ή στην ομάδα στίβου» είναι ίση με $\frac{23}{30}$, τότε ποια είναι η πιθανότητα των ενδεχομένων:

— «επιλέγουμε έναν/μία μαθητή/τρια που συμμετέχει στη θεατρική ομάδα και στην ομάδα στίβου»
— «επιλέγουμε έναν/μία μαθητή/τρια που να συμμετέχει στην θεατρική ομάδα, αλλά όχι στην ομάδα στίβου»;

## Βασικές μαθηματικές έννοιες-Ιδέες-Διεργασίες

Για τις πιθανότητες των ενδεχομένων ενός δειγματικού χώρου Ω ισχύουν οι παρακάτω ιδιότητες, γνωστές ως “κανόνες λογισμού των πιθανοτήτων”:

*Κανόνες Λογισμού Πιθανοτήτων*




$P(A΄) = 1 – P(A)$ (Π1)

Απόδειξη
Τα Α και Α΄ είναι ασυμβίβαστα ενδεχόμενα και ισχύει:

$$A \cup A΄ = Ω$$

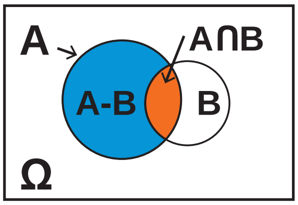
Από τον απλό προσθετικό νόμο προκύπτει ότι:

$$P(Ω) = P(A) + P(A΄) \Leftrightarrow P(A) + P(A΄) = 1 \Leftrightarrow P(A΄) = 1 - P(A)$$

ή αλλιώς: $P(A-B)=P(A)-P(A\cap B)$



***Κανόνες Λογισμού Πιθανοτήτων***




**$P(A) = P(A \cap B) + P(A - B)$ (Π2)**

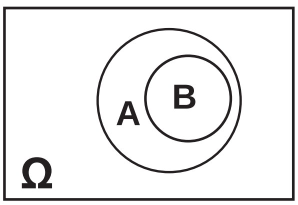
Απόδειξη

Τα ενδεχόμενα $A - B$ και $A \cap B$ είναι ασυμβίβαστα και η ένωσή τους είναι το A, δηλαδή:

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

Από τον απλό προσθετικό νόμο:

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




**Αν $B \subseteq A$ τότε $P(B) \le P(A)$ (Π3)**

Απόδειξη

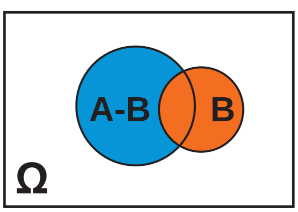
Σύμφωνα με τον (Π2): $P(A) = P(A \cap B) + P(A - B)$.

Εφόσον $B \subseteq A$, ισχύει ότι $A \cap B = B$, άρα:

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

Λαμβάνοντας υπόψη μας ότι $P(A - B) \ge 0$ προκύπτει:

$$P(A) \ge P(B)$$




**$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Π4)**

Απόδειξη

Τα ενδεχόμενα $A - B$ και B είναι ασυμβίβαστα και ισχύει:

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

Από τον απλό προσθετικό νόμο:

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

Με τη βοήθεια του (Π2) για το $A - B$:

$$P(A \cup B) = [\,P(A) - P(A \cap B)\,] + P(B) \Leftrightarrow$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Παρατηρούμε ότι:

— Για οποιαδήποτε ενδεχόμενα A και B ισχύει:

$$P(A \cup B) \le P(A) + P(B)$$

— Αν $A \cap B = \varnothing$ και συνεπώς $P(A \cap B) = 0$, τότε ο (Π4) ισοδυναμεί με τον απλό προσθετικό νόμο.

— Αν $P(A \cap B) > 0$, τότε από τον (Π4) προκύπτει η ανισότητα:

$$P(A \cup B) < P(A) + P(B)$$

## Εφαρμογές

### Εφαρμογή 1

Δίνονται τα παρακάτω ενδεχόμενα, με βάση τα δεδομένα του ερωτήματος (γ) της Διερεύνησης:

A: Ο/η μαθητής/τρια που επιλέγουμε συμμετέχει στη θεατρική ομάδα

B: Ο/η μαθητής/τρια που επιλέγουμε συμμετέχει στην ομάδα στίβου

**α)** Να υπολογίσετε την πιθανότητα του ενδεχομένου $A \cap B$.

**β)** Να υπολογίσετε την πιθανότητα του ενδεχομένου «ο/η μαθητής/τρια που επιλέγουμε συμμετέχει στην ομάδα στίβου αλλά όχι στη θεατρική ομάδα».

**γ)** Με Γ συμβολίζουμε το ενδεχόμενο «ο/η μαθητής/τρια που επιλέγουμε συμμετέχει στον όμιλο μουσικής». Υπάρχουν μαθητές/τριες του ομίλου μουσικής που συμμετέχουν σε κάποια από τις άλλες ομάδες; Αν ναι, μπορούμε να βρούμε σε ποια από τις άλλες ομάδες συμμετέχουν;

**<u>Λύση</u>**

**α)** Το ενδεχόμενο «ο/η μαθητής/τρια συμμετέχει στη θεατρική ομάδα ή στην ομάδα στίβου» είναι το $A \cup B$ άρα θα είναι $P(A \cup B) = \frac{23}{30}$.

Από τον (Π4) έχουμε $\frac{23}{30} = \frac{2}{5} + \frac{7}{15} - P(A \cap B) \Leftrightarrow P(A \cap B) = \frac{1}{10}$.

**β)** Από τον (Π2) ισχύει $P(B - A) = P(B) - P(A \cap B) = \frac{7}{15} - \frac{1}{10} = \frac{11}{30}$, που είναι η πιθανότητα του ενδεχομένου «ο/η μαθητής/τρια που επιλέγουμε συμμετέχει μόνο στην ομάδα στίβου και όχι στη θεατρική ομάδα».

**γ)** Παρατηρούμε ότι:

$$P(A \cup B) + P(\Gamma) > 1$$

Αν υποθέσουμε ότι δεν υπάρχουν μαθητές/τριες του ομίλου μουσικής που συμμετέχουν σε κάποια από τις άλλες ομάδες, τότε θα ήταν $(A \cup B) \cap \Gamma = \varnothing$. Αυτό σημαίνει ότι, λόγω του απλού προσθετικού νόμου θα έχουμε $P[(A \cup B) \cup \Gamma] = P(A \cup B) + P(\Gamma)$, άρα $P[(A \cup B) \cup \Gamma] > 1$, που είναι άτοπο. Σίγουρα, λοιπόν, η τομή $(A \cup B) \cap \Gamma$ δεν είναι κενή. Συνεπώς υπάρχουν μαθητές του ομίλου μουσικής που συμμέτεχουν και σε άλλη ομάδα (ή και στις δύο άλλες ομάδες). Ωστόσο, δεν μπορούμε να αποφανθούμε αν το Γ έχει μη κενή τομή με το Α, με το Β ή και με τα δύο. Άρα δε γνωρίζουμε σε ποια άλλη ομάδα συμμετέχουν αυτοί οι μαθητές.

## Εφαρμογή 2

Σε ένα άλλο Λύκειο από αυτό του προβλήματος της Διερεύνησης οι μαθητές/τριες έχουν τη δυνατότητα να συμμετάσχουν σε θεατρική ομάδα και ομάδα στίβου. Το 28% των μαθητών/τριών συμμετέχει στη θεατρική ομάδα, το 20% στην ομάδα στίβου και το 12% και στις δύο ομάδες. Επιλέγουμε τυχαία έναν/μία μαθητή/τρια. Ποια είναι η πιθανότητα του ενδεχομένου ο/η μαθητής/τρια που επιλέγουμε:

**α)** να συμμετέχει σε τουλάχιστον μία από τις δύο ομάδες,

**β)** να συμμετέχει μόνο στη θεατρική ομάδα,

**γ)** να συμμετέχει σε μία μόνο από τις δύο ομάδες,

**δ)** να μη συμμετέχει σε καμία ομάδα.

**Λύση**

Αν A είναι το ενδεχόμενο «ο/η μαθητής/τρια συμμετέχει στην θεατρική ομάδα» και B «ο/η μαθητής/τρια συμμετέχει στην ομάδα στίβου», τότε
$P(A) = 0,28$, $P(B) = 0,2$. Το ενδεχόμενο «ο/η μαθητής/τρια συμμετέχει και στις δύο ομάδες» είναι το $A \cap B$ και από τα δεδομένα του προβλήματος $P(A \cap B) = 0,12$.

**α)** Το ενδεχόμενο «ο/η μαθητής/τρια συμμετέχει σε τουλάχιστον μία από τις δύο ομάδες» είναι το $A \cup B$, συνεπώς από τον (Π4) με αντικατάσταση έχουμε $P(A \cup B) = 0,28 + 0,2 - 0,12 = 0,36$.

**β)** Το ενδεχόμενο «ο/η μαθητής/τρια συμμετέχει μόνο στη θεατρική ομάδα» είναι το $A - B$ και σύμφωνα με τον (Π2) ισχύει $P(A - B) = 0,28 – 0,12 = 0,16$.

**γ)** Θα χρειαστεί πρώτα να βρούμε την πιθανότητα του ενδεχομένου «ο/η μαθητής/τρια συμμετέχει μόνο στην ομάδα στίβου», δηλαδή του $B - A$. Από τον (Π2) είναι $P(B – A) = 0,2 - 0,12 = 0,08$.
Το ενδεχόμενο «ο/η μαθητής/τρια συμμετέχει σε μία μόνο από τις δύο ομάδες» είναι η ένωση των ασυμβίβαστων ενδεχομένων $A - B$ και $B - A$.
Σύμφωνα με τον απλό προσθετικό νόμο:

$$P\big((A-B)\cup(B-A)\big)=P(A-B)+P(B-A)=0{,}16+0{,}08=0{,}24$$

**δ)** Το ενδεχόμενο «ο/η μαθητής/τρια δε συμμετέχει σε καμία ομάδα» είναι το $(A \cup B)'$. Από τον (Π1) είναι:

$$P\big((A\cup B)'\big)=1-P(A\cup B)=1-0{,}36=0{,}64$$

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Το 50% των δωματίων ενός ξενοδοχείου έχουν τζάκι, το 20% έχουν καλοριφέρ και το 10% και τζάκι και καλοριφέρ. Επιλέγουμε τυχαία ένα δωμάτιο του ξενοδοχείου.
Ποια είναι η πιθανότητα του ενδεχομένου το δωμάτιο που επιλέξαμε:
**α)** να μην έχει τζάκι,
**β)** να μην έχει ούτε τζάκι ούτε καλοριφέρ,
**γ)** να έχει μόνο τζάκι;

**2)** Ας υποθέσουμε ότι A και B είναι ενδεχόμενα ενός δ.χ. Ω. Να χαρακτηρίσετε κάθε μία από τις παρακάτω προτάσεις ως σωστή ή λάθος, αιτιολογώντας την απάντησή σας:
**α)** Αν ισχύει ότι $P(A) = 0,8$ και $P(B) = 0,1$. Ισχύει ότι $B \subseteq A$, γιατί $P(B) \le P(A)$.
**β)** Αν $P(A) = 0,3$, $P(B) =0,4$ και $P(A \cup B) = 0,6$ , τότε τα A και B δεν είναι ασυμβίβαστα.

**γ)** Αν $P(A) = 0,4$ και $P(B) = 0,6$, τότε το συμπληρωματικό του A είναι το B.
**δ)** Ισχύει πάντα ότι $P(A) + P(B) \leq 1$.
**ε)** Ισχύει πάντα ότι $P(A) + P(B) - P(A \cap B) \leq 1$.
**στ)** Αν ισχύει $P(A) + P(B) = 1,5$, τότε τα A και B δεν είναι ασυμβίβαστα.
**ζ)** Αν ισχύει $P(A) + P(B) < 1$, τότε τα A και B είναι ασυμβίβαστα.
**η)** Ισχύει ότι $P(A \cap B) \leq P(A)$.

**3)** Να απαντήσετε στα ίδια ερωτήματα με την εφαρμογή 2, αν αντί για τα ποσοστά που δίνονται, γνωρίζετε αυτή τη φορά ότι το Λύκειο έχει συνολικά 120 μαθητές/τριες, από τους/τις οποίους/ες οι 32 συμμετέχουν στη θεατρική ομάδα, οι 28 στην ομάδα στίβου και 16 μαθητές/τριες συμμετέχουν και στις δύο ομάδες.

**4)** Από τους/τις μαθητές/τριες της Β΄ τάξης ενός Λυκείου το 55% είναι μαθήτριες, το 40% παίζουν μπάσκετ και το 10% είναι μαθήτριες που παίζουν μπάσκετ. Επιλέγουμε τυχαία έναν/μία μαθητή/τρια.
Να υπολογίσετε τις πιθανότητες να είναι:
**α)** μαθήτρια ή να παίζει μπάσκετ,
**β)** μαθήτρια και να μην παίζει μπάσκετ,
**γ)** μαθητής και να παίζει μπάσκετ,
**δ)** μαθητής ή να παίζει μπάσκετ.

**5)** Όλοι οι κάτοικοι μιας μικρής επαρχιακής πόλης έχουν συμβόλαιο κινητού τηλεφώνου. Το 47% των κατοίκων έχει συμβόλαιο με την εταιρεία FONATEL, το 35% των κατοίκων έχει συμβόλαιο με την TELEVIBE. Παίρνουμε τυχαία τηλέφωνο έναν κάτοικο της πόλης. Γνωρίζουμε ότι η πιθανότητα του ενδεχομένου «ο κάτοικος που πήραμε τηλέφωνο δεν έχει συμβόλαιο με καμία από τις FONATEL και TELEVIBE» είναι 23%.
Ποια είναι η πιθανότητα του ενδεχομένου, ο κάτοικος που πήραμε τηλέφωνο:
**α)** να έχει συμβόλαιο με την FONATEL ή με την TELEVIBE,
**β)** να έχει συμβόλαιο και με τις δύο εταιρείες.

**6)** Από τον πληθυσμό μιας πόλης το 42% δεν έχουν κάνει ποτέ σκι το 58% δεν έχουν ταξιδέψει ποτέ με αεροπλάνο, αλλά το 29% έχουν ήδη κάνει σκι και έχουν ταξιδέψει με αεροπλάνο. Αν πάρουμε τυχαία έναν κάτοικο της πόλης ποια είναι η πιθανότητα να μην έχει κάνει ποτε σκι και να μην έχει ταξιδέψει ποτέ με αεροπλάνο;

## Πρόσθετο Υλικό

Προτεινόμενες ασκήσεις από την παράγραφο 1.2 του βιβλίου «Άλγεβρα και Στοιχεία Πιθανοτήτων» της Α΄ Λυκείου:

Α΄ ομάδας: 7, 10, 12, 13, 14.

Β΄ ομάδας: 1, 2.

Ερωτήσεις κατανόησης 1ου κεφαλαίου: 4, 5, 10.

# ΕΝΟΤΗΤΑ 1.4. ΣΥΝΔΥΑΣΤΙΚΗ & ΠΙΘΑΝΟΤΗΤΕΣ

## Μέρος Α΄: Διατάξεις – Μεταθέσεις

### Διερεύνηση

*Ένας μεγάλος δειγματικός χώρος*

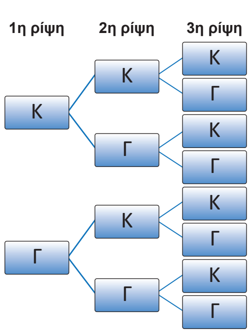
Σε ένα συγκεκριμένο μαιευτήριο κρατείται αρχείο γεννήσεων και, όπως είναι αναμενόμενο, καταγράφεται το φύλο κάθε νεογέννητου. Ποιο από τα δύο παρακάτω ενδεχόμενα είναι πιθανότερο;

— τα 2 πρώτα παιδιά του χρόνου είναι αγόρια.
— τα 5 πρώτα παιδιά του χρόνου είναι αγόρια.

### Βασικές μαθηματικές έννοιες-Ιδέες-Διεργασίες

*Η Βασική Αρχή Απαρίθμησης*

Σε πειράματα τύχης που οι δυνατές εκβάσεις είναι ισοπίθανες, ο υπολογισμός της πιθανότητας ενός ενδεχομένου ανάγεται στο μέτρημα του πλήθους των εκβάσεων, ευνοϊκών ή/και δυνατών. Συχνά, το πλήθος αυτό είναι πολύ μεγάλος αριθμός και είναι δύσκολο να βρεθεί με καταγραφή και καταμέτρηση των εκβάσεων, μία προς μία. Έτσι, σε αυτήν την παράγραφο θα μάθουμε τρόπους να μετράμε αποτελεσματικά το πλήθος των στοιχείων ενός συνόλου και θα χρησιμοποιήσουμε αυτούς τους τρόπους για να λύσουμε καθημερινά προβλήματα.




Το πλήθος των δυνατών αποτελεσμάτων τριών ρίψεων ενός αμερόληπτου κέρματος είναι ίσο με 8, όπως βλέπουμε με τη βοήθεια του διπλανού δενδροδιαγράμματος. Όμως, αν είχαμε περισσότερες ρίψεις, το δενδροδιάγραμμα θα ήταν μεγαλύτερο και η κατασκευή του δυσκολότερη. Άρα, θα ήταν σκόπιμο να είχαμε άλλον τρόπο υπολογισμού του πλήθους των δυνατών αποτελεσμάτων, χωρίς να είναι απαραίτητη η καταγραφή τους.

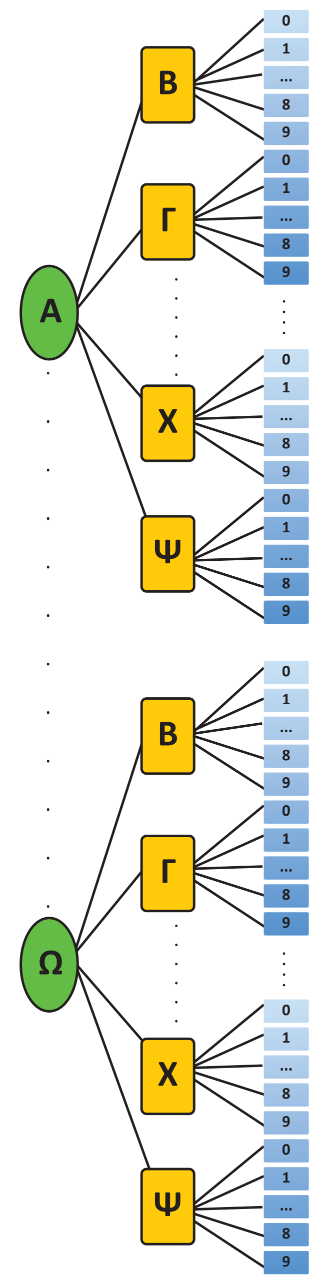
Ξεκινώντας από την περίπτωση των τριών ρίψεων, το πλήθος των δυνατών αποτελεσμάτων υπολογίζεται και με τον εξής τρόπο:

Τα δυνατά αποτελέσματα της 1ης αλλά και κάθε ρίψης είναι 2, δηλαδή Κ (κεφαλή) ή Γ (γράμματα). Για κάθε δυνατό αποτέλεσμα της 1ης ρίψης, έχουμε 2 δυνατά αποτελέσματα της 2ης ρίψης. Συνεπώς, τα δυνατά αποτελέσματα των δύο πρώτων ρίψεων είναι $2 \cdot 2 = 4$. Για κάθε ένα από αυτά έχουμε 2 δυνατά αποτελέσματα της 3ης ρίψης.

Άρα τα δυνατά αποτελέσματα των τριών ρίψεων είναι $2 \cdot 2 \cdot 2 = 2^3 = 8$.

Σκεπτόμενοι με τον ίδιο τρόπο βρίσκουμε ότι το πλήθος των δυνατών απότελεσμάτων 10 ρίψεων ενός κέρματος είναι ίσο με $2^{10}$ .

Ας δούμε ένα κάπως διαφορετικό παράδειγμα. Οι διοργανωτές ενός αγώνα δρόμου 5 χιλιομέτρων χρειάζεται να δώσουν πριν τον αγώνα σε κάθε δρομέα έναν μοναδικό κωδικό, ώστε ένα αυτοματοποιημένο σύστημα να καταγράφει τον χρόνο τερματισμού του. Ένας από τους υπεύθυνους του αγώνα πρότεινε να δοθεί σε κάθε συμμετέχοντα ένας κωδικός από τρία διατεταγμένα σύμβολα, ώστε: Το πρώτο να είναι ένα φωνήεν, το δεύτερο να είναι ένα σύμφωνο από το ελληνικό αλφάβητο και το τρίτο να είναι ένα ψηφίο του δεκαδικού συστήματος. Μία τέτοια τριάδα, για παράδειγμα, είναι ΕΜ2. Πόσοι είναι οι δυνατοί κωδικοί που μπορούν να δοθούν στους δρομείς με αυτόν τον τρόπο;




Για να απαντήσουμε, χωρίζουμε την επιλογή του κωδικού ενός δρομέα στα 5 χιλιόμετρα σε τρεις φάσεις. Στην πρώτη φάση επιλέγουμε το φωνήεν. Αυτό γίνεται με 7 τρόπους. Στη δεύτερη φάση επιλέγουμε το σύμφωνο. Τα σύμφωνα είναι 17, άρα για κάθε επιλογή φωνήεντος μπορούμε να επιλέξουμε με 17 τρόπους το σύμφωνο. Άρα τα δύο γράμματα τα επιλέγουμε με $7 \cdot 17 = 119$ διαφορετικούς τρόπους. Για κάθε έναν από αυτούς, στην τρίτη φάση επιλέγουμε έναν από τους 10 μονοψήφιους αριθμούς. Συνεπώς, οι δυνατές τριάδες είναι $119 \cdot 10 = 1190$.

Στα παραπάνω παραδείγματα εφαρμόσαμε τη **βασική αρχή απαρίθμησης**, την οποία γενικεύοντας τον προηγούμενο συλλογισμό, διατυπώνουμε ως εξής:

| Βασική Αρχή Απαρίθμησης |
|---|
| Έστω ότι μια διαδικασία μπορεί να πραγματοποιηθεί σε ν διαδοχικές φάσεις (ή επιλογές) $\varphi_1, \varphi_2, \ldots, \varphi_\nu$. Αν η $\varphi_1$ μπορεί να πραγματοποιηθεί με $\kappa_1$ τρόπους και για καθέναν από αυτούς η $\varphi_2$ μπορεί να πραγματοποιηθεί με $\kappa_2$ τρόπους,…, και για καθέναν από όλους αυτούς τους τρόπους η $\varphi_\nu$ μπορεί να πραγματοποιηθεί με $\kappa_\nu$ τρόπους, τότε η διαδικασία αυτή μπορεί να πραγματοποιηθεί με $\kappa_1 \cdot \kappa_2 \cdot \ldots \cdot \kappa_\nu$ τρόπους. |

Επιστρέφοντας στο παράδειγμα των 10 ρίψεων του κέρματος, η διαδικασία πραγματοποιείται σε 10 φάσεις (όσες είναι οι ρίψεις). Κάθε φάση πραγματοποιείται με 2 τρόπους (Κ ή Γ). Ένα δυνατό αποτέλεσμα του πειράματος τύχης, για παράδειγμα, είναι το:

(Κ, Κ, Γ, Κ, Γ, Κ, Κ, Κ, Κ, Γ)

Κάθε δυνατό αποτέλεσμα είναι μία διατεταγμένη 10-άδα από τα στοιχεία του συνόλου {Κ, Γ}. Κάθε τέτοια 10-άδα ονομάζεται **διάταξη των 2 ανά 10**

***Διατάξεις***

*Διάταξη των ν στοιχείων ενός συνόλου Α ανά κ, λέγεται καθένας από τους διαφορετικούς τρόπους με τους οποίους μπορούμε να τοποθετήσουμε στοιχεία του Α σε μια σειρά κ θέσεων επιτρέποντας επαναλήψεις.*

Το πλήθος των διατάξεων των ν ανά κ είναι ίσο με $ν^κ$.

**ή σε δεκάδες** (επιτρέποντας επαναλήψεις των στοιχείων του {Κ, Γ}). Το πλήθος των διατάξεων των 2 ανά 10 είναι ίσο με $2^{10}$.

***Διατάξεις χωρίς επαναλήψεις***

*Διάταξη των ν στοιχείων ενός συνόλου Α ανά κ χωρίς επανάληψη, με κ ≤ ν, λέγεται καθένας από τους διαφορετικούς τρόπους με τους οποίους μπορούμε να πάρουμε κ στοιχεία του Α και να τα βάλουμε σε σειρά.*

Το πλήθος των διατάξεων των ν ανά κ χωρίς επανάληψεις με κ ≤ ν συμβολίζεται με $(ν)_κ$ και είναι ίσο με $ν·(ν - 1)·...·(ν - κ + 1)$.

Ας υποθέσουμε ότι το πενταμελές μαθητικό συμβούλιο της τάξης σας συνεδριάζει για να εκλέξει 3 διαφορετικά άτομα ως πρόεδρο, γραμματέα και ταμία. Με πόσους διαφορετικούς τρόπους μπορεί να γίνει αυτή η εκλογή; Σε αυτή την περίπτωση δεν μπορεί ένας μαθητής να εκλεγεί σε πάνω από μία θέση (π.χ. πρόεδρος και γραμματέας), άρα το πλήθος των διαφορετικών τρόπων είναι μικρότερο από το πλήθος των διατάξεων των 5 ανά 3.

Η διαδικασία εκλογής μπορεί να χωριστεί σε τρεις φάσεις: 1η φάση η εκλογή προέδρου, 2η φάση η εκλογή γραμματέα και 3η φάση η εκλογή ταμία. Η 1η φάση μπορεί να γίνει με 5 τρόπους, όσα είναι και τα μέλη του πενταμελούς. Η 2η φάση μπορεί να γίνει με 4 τρόπους, όσα είναι και τα μέλη του πενταμελούς που απέμειναν ύστερα από την εκλογή του προέδρου. Η 3η φάση μπορεί να γίνει με 3 τρόπους, όσα είναι και τα μέλη του πενταμελούς που απέμειναν ύστερα και από την εκλογή του γραμματέα. Επομένως, σύμφωνα με τη βασική αρχή απαρίθμησης, το πλήθος των διαφορετικών δυνατών τριάδων είναι $5·4·3 = 60$. Κάθε τέτοια τριάδα ονομάζεται **διάταξη των 5 ανά 3 χωρίς επαναλήψεις**.

***Μεταθέσεις***

*Μετάθεση των V στοιχείων ενός συνόλου Α, λέγεται καθένας από τους διαφορετικούς τρόπους με τους οποίους μπορούμε να τα βάλουμε σε σειρά.*

Το πλήθος των μεταθέσεων ν στοιχείων είναι $ν · (ν - 1) · ... · 2 · 1$

Μία ειδική περίπτωση διατάξεων χωρίς επαναλήψεις είναι η περίπτωση που θέλουμε να βάλουμε σε σειρά τα στοιχεία ενός συνόλου και περιγράφεται στο επόμενο παράδειγμα. Ας υποθέσουμε ότι θέλουμε να βρούμε το πλήθος των διαφορετικών τρόπων που μπορούν να καθίσουν 5 θεατές μίας θεατρικής παράστασης, που μόλις μπήκαν στο θέατρο, στις 5 τελευταίες κενές θέσεις. Σύμφωνα με τη βασική αρχή απαρίθμησης οι διαφορετικοί τρόποι είναι $5·4·3·2·1 = 120$, καθώς ο πρώτος που θα καθίσει έχει 5 επιλογές, ο δεύτερος έχει 1 λιγότερη επιλογή, δηλαδή 4, ο τρίτος έχει 3 επιλογές, ο τέταρτος 2 επιλογές και ο πέμπτος 1 επιλογή. Κάθε διαφορετικός τρόπος να καθίσουν είναι μία **μετάθεση** 5 στοιχείων.

***Ένας χρήσιμος συμβολισμός: Το παραγοντικό***

Όπως είδαμε (και θα φανεί ακόμα περισσότερο στη συνέχεια) είναι χρήσιμο να υπολογίζουμε γινόμενα διαδοχικών φυσικών αριθμών. Συχνά θα χρειαστεί να υπολογιστούν τέτοια μεγάλα γινόμενα. Το γινόμενο των φυσικών αριθμών 1, 2, 3,…..,ν-1, ν συμβολίζεται ως ν! και διαβάζεται «ν παραγοντικό».

Εξ ορισμού $0! = 1$

Είναι $ν! =1·2· 3·.....· (ν-1) ·ν$ . Για παράδειγμα, $5! =5·4·3·2·1$ .

Με τη χρήση του παραγοντικού το πλήθος των μεταθέσεων 5 στοιχείων γράφεται 5!, ενώ το πλήθος των διατάξεων των 5 ανα 3 χωρίς επαναλήψεις, δηλαδή 5·4·3 γράφεται και ως $\frac{5!}{(5-3)!}$, που σε κάποιες περιπτώσεις είναι ένας «βολικός» συμβολισμός.

— *Το πλήθος των μεταθέσεων ν στοιχείων είναι ν!*

— *Με αυτόν τον συμβολισμό, το πλήθος των διατάξεων των ν ανά κ χωρίς επαναλήψεις είναι*

$$(\nu)_\kappa = \frac{\nu!}{(\nu-\kappa)!}$$

## Εφαρμογές

### Εφαρμογή 1

Σε μία τάξη υπάρχουν ακριβώς 25 καρέκλες. Με πόσους διαφορετικούς τρόπους μπορούν να καθίσουν σε αυτές οι 25 μαθητές/τριες μιας τάξης;

**Λύση**

Κάθε τρόπος είναι μία μετάθεση των 25 μαθητών/τριών. Άρα οι διαφορετικοί τρόποι να καθίσουν είναι πλήθους 25!, δηλαδή γύρω στα 15,5 πεντάκις εκατομμύρια.

### Εφαρμογή 2

Ο ιδιοκτήτης ενός ηλεκτρονικού υπολογιστή έχει ξεχάσει τον προσωπικό του κωδικό χρήστη για να εισέλθει στον λογαριασμό του. Το μόνο που θυμάται είναι ότι ο κωδικός αποτελείται από οκτώ πεζά γράμματα του αγγλικού αλφαβήτου. Σκέφτεται να ετοιμάσει μια λίστα με όλους τους πιθανούς κωδικούς. Ο χρόνος που χρειάζεται για να γράψει σε χαρτί έναν τυχαίο τέτοιο κωδικό χρήστη είναι 3 δευτερόλεπτα.

α) Πόσο χρόνο χρειάζεται ο χρήστης για γράψει τη λίστα;

β) Ο χρήστης θυμήθηκε ότι στον οκταψήφιο κωδικό του, κανένα γράμμα δεν επαναλαμβάνεται. Πόσος χρόνος χρειάζεται για να γραφτεί η λίστα, σε αυτή την περίπτωση;

**Λύση**

Το αγγλικό αλφάβητο έχει 26 γράμματα.

α) Κάθε πιθανός κωδικός είναι μία διάταξη των 26 γραμμάτων σε οκτάδα, με δυνατές τις επαναλήψεις, άρα το πλήθος των πιθανών κωδικών είναι $26^8$ = 208.827.064.576. Συνεπώς, ο απαιτούμενος χρόνος για να γραφτεί η λίστα είναι 3·208.827.064.576 δευτερόλεπτα. Επίσης, κάθε εικοσιτετράωρο έχει 60·60·24 = 86400 δευτερόλεπτα. Άρα για να γραφτεί η λίστα χρειάζονται 7.250.940 εικοσιτετράωρα.

β) Σε αυτή την περίπτωση κάθε κωδικός είναι μία διάταξη των 26 γραμμάτων σε οκτάδα, χωρίς να επιτρέπονται επαναλήψεις γραμμάτων, άρα το πλήθος των πιθανών κωδικών είναι $\frac{26!}{(26-8)!}=$

$$\frac{26!}{18!}=\frac{26\cdot 25\cdot \ldots \cdot 19\cdot 18!}{18!}=26\cdot 25\cdot \ldots \cdot 19=62.990.928.000$$

Ο χρόνος που χρειάζεται να γραφτεί η λίστα είναι 188.972.784.000 δευτερόλεπτα, ή 2.187.185 εικοσιτετράωρα.

## Εφαρμογή 3

Μία τράπουλα έχει 52 διαφορετικά φύλλα. Αν ανακατέψουμε την τράπουλα και ανοίξουμε τα 4 πρώτα φύλλα, ποια είναι η πιθανότητα αυτά να είναι άσοι;

### Λύση

Πρώτα θα υπολογίσουμε με πόσους διαφορετικούς τρόπους μπορούν να μπούν σε σειρά (δηλαδή να μετατεθούν) τα 52 φύλλα μετά το ανακάτεμα. Πρόκειται για μεταθέσεις 52 στοιχείων, άρα το πλήθος είναι 52!

Στη συνέχεια, ας υποθέσουμε ότι τα 4 πρώτα φύλλα είναι άσοι. Πρώτα υπολογίζουμε το πλήθος των διαφορετικών τρόπων που έχουν μετατεθεί τα υπόλοιπα 48 φύλλα. Αυτό είναι ίσο με 48!.

Όμως για κάθε έναν από αυτούς τους 48! τρόπους, υπάρχουν 4! τρόποι να εμφανιστούν οι 4 άσοι, που -όπως υποθέσαμε- είναι πρώτοι, πριν τα 48 φύλλα. Συνεπώς, από τη βασική αρχή απαρίθμησης το πλήθος των διαφορετικών τρόπων να εμφανιστούν πρώτα οι 4 άσοι είναι:

$$(\text{μεταθέσεις 4 στοιχείων})\cdot(\text{μεταθέσεις 48 στοιχείων}) = 4!\cdot 48!$$

Συνεπώς, αν θεωρήσουμε το πείραμα τύχης «ανακατεύουμε μία συνηθισμένη τράπουλα 52 φύλλων» το πλήθος των δυνατών αποτελεσμάτων (δηλαδή των δυνατών μεταθέσεων των φύλλων της τράπουλας) είναι 52!. Θεωρούμε επίσης ότι όλα τα δυνατά αποτελέσματα είναι ισοπίθανα. Έστω το ενδεχόμενο Α: «ανοίγουμε τα 4 πρώτα φύλλα της τράπουλας και αυτά είναι άσοι», τότε το πλήθος των ευνοϊκών αποτελεσμάτων είναι 4!·48.

Άρα από τον κλασικό ορισμό

$$P(A) = \frac{4!\cdot 48!}{52!} = \frac{4!\cdot 48!}{52\cdot 51\cdot 50\cdot 49\cdot 48!} = \frac{4!}{52\cdot 51\cdot 50\cdot 49} = \frac{6}{1.624.350} = \frac{1}{270.725}$$

## Εφαρμογή 4

Δύο φοιτητές, ο Βαγγέλης και η Μαρία θέλουν να ταξιδέψουν με το λεωφορείο που εκτελεί το δρομολόγιο των 5:30 «Αθήνα-Πάτρα», αλλά υπάρχουν μόνο 4 κενές θέσεις, οι 2, 13, 14 και 56. Για τις 4 αυτές θέσεις υπάρχουν 7 υποψήφιοι επιβάτες (μαζί με το Βαγγέλη και τη Μαρία). Όμως κανείς τους δεν προηγείται· θα γίνει κλήρωση μεταξύ τους για το ποιος θα ταξιδέψει με αυτό το δρομολόγιο. Να βρείτε την πιθανότητα να ταξιδέψουν και τα δύο παιδιά:

**α)** ο Βαγγέλης στη θέση 56 κι η Μαρία στη θέση 2,

**β)** σε διπλανές θέσεις.


***Κάτοψη του λεωφορείου***

### Λύση

α) Οι τρόποι που μπορούν να κληρωθούν 4 θέσεις σε 7 επιβάτες είναι 7·6·5·4 = 840 (διατάξεις των 7 ανά 4 χωρίς επαναλήψεις). Αν θεωρήσουμε το πείραμα τύχης «κληρώνουμε τις 4 θέσεις στους 7 επιβάτες» και θεωρήσουμε ισοπίθανες όλες τις πιθανές εκβάσεις, τότε αυτές είναι 840.

Πρέπει να υπολογίσουμε το πλήθος ευνοϊκών εκβάσεων του ενδεχομένου «ο Βαγγέλης στη θέση 56 κι η Μαρία στη θέση 2». Όπως φαίνεται στο διπλανό σχήμα αρκεί να βρούμε με πόσους τρόπους μπορούν να συμπληρωθούν οι θέσεις 13 και 14, δηλαδή οι 2 θέσεις που μένουν για τα υπόλοιπα 5 άτομα (διατάξεις των 5 ανά 2 χωρίς επαναλήψεις). Αυτό γίνεται με $5 \cdot 4 = 20$ τρόπους. Συνεπώς, από τον κλασικό ορισμό, η πιθανότητα να καθίσουν ο Βαγγέλης και η Μαρία στις θέσεις 56 και 2 είναι $\frac{20}{840} = \frac{1}{42}$.

| *Θέση 56* | *Θέση 2* | *Θέση 13* | *Θέση 14* |
|---|---|---|---|
| *Βαγγ* | *Μαρ* | *;* | *;* |

*Για το ερώτημα (α)*

**β)** Οι μόνες διπλανές θέσεις είναι οι 13 και 14. Οι ευνοϊκές εκβάσεις είναι εκείνες που ο Βαγγέλης και η Μαρία κληρώνονται στις θέσεις αυτές, χωρίς να έχει σημασία ποιος από τους δύο κληρώνεται στην κάθε θέση. Υποθέτουμε ότι ο Βαγγέλης και η Μαρία κληρώνονται σε αυτές τις θέσεις. Ομοίως με το (α), οι υπόλοιποι επιβάτες κληρώνονται με 20 τρόπους στις θέσεις που μένουν (56 και 2). Ωστόσο, για κάθε έναν από αυτούς τους τρόπους υπάρχουν 2 μεταθέσεις του Βαγγέλη και της Μαρίας στις θέσεις 13 και 14. Άρα από την βασική αρχή απαρίθμησης οι ευνοϊκές εκβάσεις είναι $2 \cdot 20 = 40$. Άρα η πιθανότητα ο Βαγγέλης και η Μαρία να καθίσουν σε διπλανές θέσεις είναι ίση με $\frac{40}{840} = \frac{1}{21}$.

| *Θέση 13* | *Θέση 14* | *Θέση 56* | *Θέση 2* |
|---|---|---|---|
| *Βαγγ* | *Μαρ* | *;* | *;* |
| *Μαρ* | *Βαγγ* | | |

*Για το ερώτημα (β)*

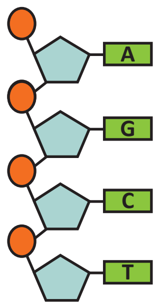
## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Οι διοργανωτές αγώνων δρόμου 5 και 10 χιλιομέτρων, που θα γίνουν την ίδια ημέρα, θέλουν να δώσουν έναν μοναδικό κωδικό σε κάθε συμμετέχοντα. Επίσης, θα ήθελαν κάθε κωδικός συμμετέχοντα στα 5 χιλιόμετρα να είναι ευδιάκριτος από τον κωδικό ενός συμμετέχοντα στα 10 χιλιόμετρα. Στα 5 χιλιόμετρα υπάρχουν 2.573 συμμετέχοντες, ενώ στα 10 χιλιόμετρα υπάρχουν 1.113 συμμετέχοντες. Αποφασίστηκε για τα 5 χιλιόμετρα να δοθεί σε κάθε συμμετέχοντα ένας κωδικός από τρία διατεταγμένα σύμβολα, ώστε:
Το πρώτο και το δεύτερο να είναι τυχαίο σύμφωνο από το ελληνικό αλφάβητο και το τρίτο να είναι ένας τυχαίος μονοψήφιος αριθμός, π.χ. ΝΜ2.

**α)** Επαρκούν οι διατεταγμένες τριάδες αυτές για όλους τους συμμετέχοντες στα 5 χιλιόμετρα;

**β)** Τι θα προτείνατε στους διοργανωτές να κάνουν για τα 10 χιλιόμετρα;

**2)** Με πόσους τρόπους μπορούν να τοποθετηθούν τα γράμματα Α, Ε, Ε, Θ, Ι, Υ το ένα μετά από το άλλο; Ποια είναι η πιθανότητα, αν τοποθετήσουμε τα γράμματα σε τυχαία σειρά, να σχηματιστεί η λέξη ΕΥΘΕΙΑ;




**3)** Οι αζωτούχες βάσεις που μπορεί να έχει ένα νουκλεοτίδιο είναι η Αδενίνη (A), η Γουανίνη (G), η Κυτοσίνη (C) και η Θυμίνη (T). Τα νουκλεοτίδια, ανάλογα με την σειρά τους σε τριάδες, καθορίζουν ποιο αμινοξύ θα τοποθετηθεί στην αντίστοιχη θέση κατά τη σύνθεση των πρωτεϊνών. Οι τριάδες αυτές ονομάζονται κωδικόνια.

**α)** Πόσα διαφορετικά κωδικόνια μπορούν να σχηματιστούν;

**β)** Συνολικά τα κωδικόνια αντιστοιχούν στον σχηματισμό 20 αμινοξέων. Αν επιλέξουμε τυχαία ένα κωδικόνιο, είναι ίδια η πιθανότητα να αντιστοιχεί σε ένα από τα 20 αμινοξέα (θα χρειαστεί να αναζητήσετε πληροφορίες για την αντιστοιχία κωδικονίων και αμινοξέων για να απαντήσετε);

**4)** Αν κάποιος διαθέτει 2 μπουφάν (ένα μαύρο κι ένα μπλε), 4 παντελόνια, 3 μπλούζες, 10 ζευγάρια κάλτσες και 3 ζευγάρια παπούτσια, με πόσους τρόπους μπορεί να ντυθεί, φορώντας από όλα τα είδη; Αν ένας επιλέξει τυχαία έναν από αυτούς τους τρόπους για να ντυθεί φορώντας ένα μπουφάν, ένα παντελόνι, μία μπλούζα, ένα ζευγάρι κάλτσες και ένα ζευγάρι παπούτσια, ποια είναι η πιθανότητα να φοράει το μπλε μπουφάν;

**5)** Αν ρίξουμε ένα αμερόληπτο ζάρι 4 φορές ποια είναι η πιθανότητα να φέρουμε 4 διαφορετικά αποτελέσματα;

**6)** **α)** Πόσες πινακίδες κυκλοφορίας μπορούμε να κατασκευάσουμε που περιέχουν στη σειρά τρία κεφαλαία γράμματα του ελληνικού αλφαβήτου ακολουθούμενα από έναν τετραψήφιο αριθμό;

**β)** Από το σύνολο των πινακίδων που περιγράφονται στο ερώτημα (α), μπορούν να χρησιμοποιηθούν εκείνες που τα γράμματά τους ανήκουν και στο λατινικό αλφάβητο. Αν επιλέξουμε τυχαία μία πινακίδα του ερωτήματος, (α) ποια είναι η πιθανότητα να είναι κατάλληλη προς χρήση;

**7)** Ο Θανάσης, ο Μιχάλης, ο Κώστας, ο Αντρέι κι ο Δημήτρης είναι οι παίκτες της σχολικής ομάδας μπάσκετ των αγοριών του Γ1 και ο προπονητής της ομάδας πρόκειται να τους δώσει τις εμφανίσεις τους για τους σχολικούς αγώνες.

**α)** Οι διαθέσιμες εμφανίσεις έχουν τυπωμένα τα νούμερα 7, 13, 15, 20 και 27. Με πόσους τρόπους μπορούν να μοιραστούν οι εμφανίσεις στους μαθητές; Αν οι εμφανίσεις μοιραστούν τυχαία στους παίκτες, ποια είναι η πιθανότητα να πάρει ο Δημήτρης τη φανέλα με το 7; Ποια είναι η πιθανότητα να πάρει ο Δημήτρης τη φανέλα με το 7 και να μην πάρει ο Θανάσης τη φανέλα με το 13;

**β)** Να απαντήσετε στα ίδια ερωτήματα, αν οι διαθέσιμες εμφανίσεις είναι οι 7, 11, 13, 15, 19, 20 και 27.

**8)** Με πόσους διαφορετικούς τρόπους μπορούν να καθίσουν 4 άτομα σε 6 θέσεις μιας σειράς; Ποια είναι η πιθανότητα η τελευταία θέση να μείνει κενή;

**9)** Με πόσους διαφορετικούς τρόπους μπορούν να μπουν σε μια σειρά 4 αγόρια και 3 κορίτσια; Αν επιλέξουμε τη σειρά των 7 παιδιών τυχαία, ποια είναι η πιθανότητα να είναι όλα μαζί τα αγόρια και όλα μαζί τα κορίτσια;

**10)** Δέκα παιδιά, μεταξύ των οποίων ο Κώστας και η Ελένη, θα καθίσουν τυχαία ο ένας δίπλα στον άλλον σε δέκα θέσεις. Ποια είναι η πιθανότητα ο Κώστας και η Ελένη να καθίσουν σε διπλανές θέσεις;

**11)** Να βρείτε την πιθανότητα του ενδεχομένου τέσσερα τυχαία επιλεγμένα άτομα από την τάξη σας (ή το σχολείο σας), να έχουν γεννηθεί σε τέσσερις διαφορετικές εποχές του έτους.

# Μέρος Β΄: Συνδυασμοί

## Διερεύνηση

Έχετε να τοποθετήσετε τρεις επιστολές σε φακέλους. Επίσης έχετε στη διάθεσή σας τέσσερις φακέλους διαφορετικού χρώματος: κίτρινο, μπλε, κόκκινο και πράσινο. Μόνο μία επιστολή μπαίνει σε κάθε φάκελο. Αν κάνετε τυχαία την επιλογή των φακέλων που θα χρησιμοποιήσετε, πόσοι τρόποι υπάρχουν να τοποθετηθούν οι επιστολές στους φακέλους:

**α)** αν η πρώτη επιστολή είναι ευχαριστήρια, η δεύτερη είναι συγχαρητήρια και η τρίτη είναι πρόσκληση;

**β)** αν και οι τρεις επιστολές είναι ακριβώς ίδιες;

## Βασικές μαθηματικές έννοιες-Ιδέες-Διεργασίες

***Συνδυασμοί***

Συχνά είναι χρήσιμο να υπολογίσουμε το πλήθος των δυνατών τρόπων να επιλεγούν κ από τα ν στοιχεία ενός συνόλου, χωρίς να μας ενδιαφέρει η σειρά επιλογής τους. Κάθε τέτοιος τρόπος ονομάζεται συνδυασμός των ν ανά κ. Με στόχο να βρούμε έναν τρόπο υπολογισμού του πλήθους των συνδυασμών των ν ανά κ, στη γενική περίπτωση, ας δούμε πρώτα το επόμενο παράδειγμα.

*Συνδυασμός των ν στοιχείων ενός συνόλου Α ανά κ λέγεται κάθε υποσύνολο του Α με κ στοιχεία.*

Ας υποθέσουμε ότι από τους 5 παίκτες μίας σχολικής ομάδας μπάσκετ πρέπει να επιλεγούν 3 για να λάβουν μέρος στο σχολικό πρωτάθλημα μπάσκετ με τίτλο «τρεις εναντίον τριών». Ο καθηγητής Φυσικής Αγωγής του σχολείου, αφού επιλέξει τους παίκτες, πρέπει να γράψει τα ονόματά τους στο διπλανό καρτελάκι δίπλα στα νούμερα που θα έχουν στις εμφανίσεις τους. Κάθε παίκτης θα έχει

| Ομάδα μπάσκετ «3 εναντίον 3» | |
|---|---|
| Νούμερο | Ονοματεπώνυμο |
| 7 | |
| 10 | |
| 19 | |

διαφορετικό νούμερο. Όπως φαίνεται οι εμφανίσεις έχουν το 7, το 10 και το 19.
Με πόσους διαφορετικούς τρόπους μπορεί να συμπληρωθεί το καρτελάκι;
Το καρτελάκι μπορεί να συμπληρωθεί με 60 διαφορετικούς τρόπους, αφού πρόκειται για διατάξεις των 5 ανά 3 χωρίς επανάληψη, άρα το πλήθος τους είναι $5{\cdot}4{\cdot}3$ ή αλλιώς $\frac{5!}{(5-3)!}$.

Για μια δεδομένη επιλογή των 3 παικτών με πόσους διαφορετικούς τρόπους μπορεί να συμπληρωθεί το καρτελάκι;
Ας υποθέσουμε ότι οι 3 παίκτες που έχουν επιλεγεί είναι ο Αντώνης, ο Βασίλης και ο Γιάννης. Το καρτελάκι μπορεί να συμπληρωθεί με τους εξής 6 τρόπους:

| Νούμερο | Αρχικό γράμμα ονόματος του παίκτη | | | | | |
|---|---|---|---|---|---|---|
| 7 | Α | Α | Β | Β | Γ | Γ |
| 10 | Β | Γ | Α | Γ | Α | Β |
| 19 | Γ | Β | Γ | Α | Β | Α |

Θα μπορούσαμε να υπολογίσουμε το πλήθος των τρόπων χωρίς την καταγραφή τους, καθώς πρόκειται για μεταθέσεις 3 στοιχείων, που το πλήθος τους είναι $3! = 6$.

Πόσες είναι οι δυνατές τριάδες παικτών που μπορούν προκύψουν από τους 5 παίκτες;
Ας υποθέσουμε ότι x είναι το πλήθος των διαφορετικών τριάδων. Και οι 6 μεταθέσεις των Α, Β και Γ του παραπάνω πίνακα αντιστοιχούν σε έναν από τους x αυτούς τρόπους, αφού οι παίκτες που επιλέγονται είναι οι ίδιοι (οι Α, Β και Γ απλώς αλλάζει το νούμερο στην εμφάνιση του κάθε ενός). Το x ισούται με το πλήθος των συνδυασμών των 5 ανά 3. Λόγω της βασικής αρχής απαρίθμησης:

$$\begin{pmatrix}\text{το πλήθος των}\\ \text{δυνατών τριάδων}\end{pmatrix}\cdot\begin{pmatrix}\text{το πλήθος των}\\ \text{διαφορετικών τρόπων}\\ \text{που μπορεί να}\\ \text{συμπληρωθεί}\\ \text{το καρτελάκι για}\\ \text{δεδομένη τριάδα}\end{pmatrix}=\begin{pmatrix}\text{το πλήθος των}\\ \text{διαφορετικών}\\ \text{τρόπων που μπορεί}\\ \text{να συμπληρωθεί}\\ \text{το καρτελάκι}\end{pmatrix}$$

Άρα:

$$x\cdot 3!=\frac{5!}{(5-3)!}\Leftrightarrow x=\frac{5!}{3!(5-3)!}\Leftrightarrow x=\frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1\cdot 2\cdot 1}\Leftrightarrow x=10$$

*Για το πλήθος των συν-δυασμών των ν ανά κ ισχύει:*

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

Κάθε τρόπο από τους 10 τον ονομάζουμε **συνδυασμό των 5 ανά 3**.
Για το πλήθος των συνδυασμών των 5 ανά 3 χρησιμοποιούμε τον συμβολισμό:

$$\binom{5}{3}=\frac{5!}{3!(5-3)!}$$

Η ισότητα $x\cdot 3!=\frac{5!}{(5-3)!}$ που σκεφτήκαμε και χρησιμοποιήσαμε για να λύσουμε το πρόβλημα, σημαίνει ότι:

$$\begin{pmatrix}\text{Πλήθος}\\ \text{συνδυασμών}\\ \text{των 5 ανά 3}\end{pmatrix}\cdot\begin{pmatrix}\text{Πλήθος}\\ \text{μεταθέσεων}\\ \text{των 3}\end{pmatrix}=\begin{pmatrix}\text{Πλήθος}\\ \text{διατάξεων}\\ \text{των 5 ανά 3}\\ \text{χωρίς επαν.}\end{pmatrix}$$

Και μπορεί να γενικευτεί:

$$\begin{pmatrix}\text{Πλήθος}\\ \text{συνδυασμών}\\ \text{των ν ανά κ}\end{pmatrix}\cdot\begin{pmatrix}\text{Πλήθος}\\ \text{μεταθέσεων}\\ \text{των κ}\end{pmatrix}=\begin{pmatrix}\text{Πλήθος}\\ \text{διατάξεων}\\ \text{των ν ανά κ}\\ \text{χωρίς επαν.}\end{pmatrix}$$

Ή αλλιώς:

$$\begin{pmatrix}\text{Πλήθος}\\ \text{συνδυασμών}\\ \text{των ν ανά κ}\end{pmatrix}=\frac{\begin{pmatrix}\text{Πλήθος}\\ \text{διατάξεων}\\ \text{των ν ανά κ}\\ \text{χωρίς επαν.}\end{pmatrix}}{\begin{pmatrix}\text{Πλήθος}\\ \text{μεταθέσεων}\\ \text{των κ}\end{pmatrix}}$$

## Εφαρμογές

### Εφαρμογή 1

Το πενταμελές συμβούλιο του σχολείου αποτελείται από 4 κορίτσια και 1 αγόρι. Πρόκειται να επιλεγούν, με κλήρωση, 3 από τα 5 μέλη του συμβουλίου για να εκπροσωπήσουν την τάξη τους σε μία ενημερωτική συνάντηση με τους καθηγητές τους. Ποια είναι η πιθανότητα και τα 3 μέλη που θα επιλεγούν να είναι κορίτσια;

**Λύση**

Θεωρούμε το πείραμα τύχης: «Επιλέγουμε τυχαία 3 από τα 5 μέλη του πενταμελούς». Η σειρά επιλογής (ποιος/α θα επιλεγεί πρώτος/η, δεύτερος/η και τρίτος/η) δεν έχει σημασία. Οι τρόποι επιλογής είναι το πλήθος των συνδυασμών των 5 ανα 3, δηλαδή $\binom{5}{3}=10$. Αυτό είναι και το πλήθος των αποτελεσμάτων του δ.χ. Ω του πειράματος τύχης (και αν η κλήρωση είναι δίκαια μπορούμε να θεωρήσουμε ότι είναι ισοπίθανα).

Θεωρούμε ως Α το ενδεχόμενο του Ω: «επιλέγονται μόνο κορίτσια». Τότε το πλήθος των ευνοϊκών αποτελεσμάτων του Α είναι το πλήθος των συνδυασμών των 4 ανά 3. Πράγματι, εξαιρώντας το αγόρι, τότε μένει να βρούμε με πόσους τρόπους επιλέγονται 3 από τα 4 κορίτσια.

Από την ισότητα $\binom{4}{3}\cdot 3!=\frac{4!}{(4-3)!}$ προκύπτει ότι:

$$\binom{4}{3}=\frac{4!}{3!(4-3)!}=\frac{4\cdot 3!}{3!\cdot 1!}=4$$

Άρα, από τον κλασικό ορισμό της πιθανότητας:

$$P(A)=\frac{4}{10}=0,4.$$

## Εφαρμογή 2

Στην εφαρμογή 4 του μέρους Α΄, υπολογίσαμε την πιθανότητα του ενδεχομένου «ο Βαγγέλης και η Μαρία κληρώνονται για να ταξιδέψουν σε διπλανές θέσεις».

**α)** Να υπολογίσετε την πιθανότητα του ενδεχομένου «ο Βαγγέλης και η Μαρία κληρώνονται να ταξιδέψουν χωρίς να είναι σε διπλανές θέσεις».

**β)** Να υπολογίσετε την πιθανότητα του ενδεχομένου «ο Βαγγέλης κληρώνεται για να ταξιδέψει».

**Λύση**

**α)** Αν Α είναι το ενδεχόμενο «ο Βαγγέλης και η Μαρία ταξιδεύουν σε διπλανές θέσεις» και Β είναι το ενδεχόμενο «ο Βαγγέλης και οι Μαρία κληρώνονται να ταξιδέψουν», τότε το Β - Α είναι το ενδεχόμενο «ο Βαγγέλης και η Μαρία ταξιδεύουν σε θέσεις που δεν είναι διπλανές». Από τον Π.4 ισχύει ότι $P(B - A) = P(B) - P(A \cap B)$.

Επίσης ισχύει ότι $A \subseteq B$, άρα $A \cap B = A$.

Άρα $P(B - A) = P(B) - P(A)$. Επομένως, έχοντας υπολογίσει ότι $P(A)=\frac{1}{21}$, αρκεί να υπολογίσουμε την P(B). Οι διαφορετικοί τρόποι που μπορούν να επιλεγούν 4 επιβάτες από τους 7 για να ταξιδέψουν, χωρίς να μας ενδιαφέρει ποιος θα ταξιδέψει σε ποια θέση είναι :

$$\binom{7}{4}=\frac{7!}{4!\cdot 3!}=\frac{7\cdot 6\cdot 5\cdot 4!}{4!\cdot 3!}=\frac{7\cdot 6\cdot 5}{3!}=35$$

Αυτό μπορεί να θεωρηθεί ως το πλήθος των στοιχείων του δ.χ. του πειράματος τύχης «επιλέγονται 4 επιβάτες για να ταξιδέψουν, χωρίς να μας ενδιαφέρει η σειρά επιλογής τους». Ομοίως, τα ευνοϊκά αποτελέσματα είναι οι τετράδες που περιέχουν τον Βαγγέλη, τη Μαρία και ακόμα 2 άτομα επιλεγμένα από τους υπόλοιπους 5 επιβάτες:

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = 10$$

Άρα:

$$P(B) = \frac{10}{35} = \frac{2}{7}$$

και η πιθανότητα να ταξιδέψουν ο Βαγγέλης και η Μαρία σε όχι διπλανές θέσεις είναι ίση με:

$$P(B-A) = \frac{2}{7} - \frac{1}{21} = \frac{5}{21}$$

**β)** Θεωρούμε ως το ενδεχόμενο Γ: «ο Βαγγέλης κληρώνεται να ταξιδέψει». Το πλήθος των ευνοϊκών εκβάσεων του Γ υπολογίζεται ως εξής:
Θεωρούμε ότι ο Βαγγέλης έχει κληρωθεί. Με πόσους τρόπους μπορούν να κληρωθούν οι υπόλοιποι 6 επιβάτες στις 3 θέσεις που απομένουν, χωρίς να μας ενδιαφέρει η σειρά κλήρωσης;
Το πλήθος αυτών των τρόπων είναι ίσο με:

$$\binom{6}{3} = \frac{6!}{3! \cdot (6-3)!} = 20$$

Τελικά:

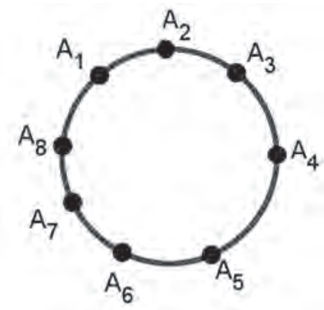
$$P(\Gamma) = \frac{20}{35} = \frac{4}{7}$$

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Σε ένα πρωτάθλημα συμμετέχουν 7 ομάδες και αγωνίζονται όλες με όλες μία φορά. Να υπολογίσετε πόσοι αγώνες θα γίνουν.

**2)** **α)** Από ένα σύνολο 10 μαθητών επιλέγουμε 4 μαθητές τυχαία. Ποια είναι η πιθανότητα να επιλέξουμε έναν συγκεκριμένο μαθητή;
**β)** Να απαντήσετε στο ίδιο ερώτημα αν έχουμε ένα σύνολο ν μαθητών και επιλέγουμε, τυχαία, κ μαθητές.

**3)** Σε ένα κουτί υπάρχουν 5 μπαταρίες, από τις οποίες οι 2 είναι αποφορτισμένες. Επιλέγουμε τυχαία 2 μπαταρίες από το κουτί. Ποια είναι η πιθανότητα του ενδεχομένου:
**α)** Οι μπαταρίες που επιλέξαμε είναι αποφορτισμένες.
**β)** Το πολύ μία από τις μπαταρίες που επιλέξαμε είναι αποφορτισμένη.
**γ)** Οι μπαταρίες που επιλέξαμε δεν είναι αποφορτισμένες.




**4)** Σε έναν κύκλο δίνονται 8 σημεία $A_1, A_2, \ldots, A_8$.
**α)** Πόσα ευθύγραμμα τμήματα ορίζουν τα σημεία αυτά;
**β)** Πόσες διαγώνιες έχει ένα κανονικό οκτάγωνο;
**γ)** Πόσα τρίγωνα υπάρχουν με αυτά τα σημεία ως κορυφές;
**δ)** Επιλέγουμε τυχαία ένα από τα ευθύγραμμα τμήματα του ερωτήματος (α). Ποια είναι η πιθανότητα:
**i)** να μη διέρχεται από το σημείο $A_1$;
**ii)** να διέρχεται από το σημείο $A_2$;

**5)** Από ένα σύλλογο καθηγητών με 7 άνδρες και 6 γυναίκες επιλέγουμε τυχαία 4 άτομα. Να βρείτε τις πιθανότητες των ενδεχομένων:
**α)** τα άτομα να είναι γυναίκες,
**β)** ένα τουλάχιστον να είναι άνδρας,
**γ)** να υπάρχει μία μόνο γυναίκα.

**6)** Ένα κουτί περιέχει 20 ηλεκτρικές ασφάλειες, από τις οποίες οι 5 είναι ελαττωματικές. Από ένα σύστημα ελέγχου ποιότητας επιλέγονται τυχαία 4 ασφάλειες και δοκιμάζονται.
Αν βρεθούν περισσότερες από μία ελαττωματικές, το κουτί επιστρέφεται ως απαράδεκτο. Να βρείτε την πιθανότητα να επιστραφεί ως απαράδεκτο ένα κουτί που έχει 5 ελαττωματικές ασφάλειες.

**7)** Έχετε δύο σύμβολα το X και I. Πόσες διαφορετικές συμβολοσειρές μπορείτε να σχηματίσετε χρησιμοποιώντας 4 φορές το I και 3 φορές το X; Π.χ. μία συμβολοσειρά είναι η XIIXIIX. Αν κάποιος επιλέξει τυχαία μία τέτοια σειρά, ποια είναι η πιθανότητα το 1ο σύμβολο από αριστερά να είναι X;

## Πρόσθετο Υλικό

**1)** Τι από τα παρακάτω είναι πιθανότερο;
— Να κερδίσετε στο τζόκερ έχοντας συμπληρώσει μία στήλη.
— Να καλέσετε στο τηλέφωνο έναν φίλο ή μία φίλη σας, γνωρίζοντας μόνο τρία από τα δέκα νούμερα του τηλεφώνου του και επιλέγοντας τα υπόλοιπα επτά νούμερα στην τύχη.

**2)** **α)** Εικοσιτρείς φοιτητές/τριες έχουν μια κοινή ομάδα συζήτησης, σε δικτυακή εφαρμογή επικοινωνίας, για ανταλλαγή σημειώσεων. Σε αυτή την ομάδα ξεκινάει μία συζήτηση για το πότε έχει καθένας/καθεμία γενέθλια. Ποια είναι η πιθανότητα να έχουν δύο τουλάχιστον φοιτητές/τριες γενέθλια την ίδια ημέρα;
**β)** Να εκτιμήσετε το πλήθος των ατόμων μίας ομάδας, ώστε να είναι περισσότερο από 99,9% πιθανό ότι θα υπάρχουν τουλάχιστον δύο άτομα που έχουν γενέθλια την ίδια ημέρα.

Η τιμή της παράστασης $1-\dfrac{365\cdot 364\cdot \ldots \cdot(365-\kappa+1)}{365^{\kappa}}$, όπου κ ≤ 365 φυσικός αριθμός δίνεται (προσεγγιστικά) στον παρακάτω πίνακα, για κάποιες τιμές του κ:

| **κ** | **10** | **20** | **23** | **30** | **60** | **70** |
|---|---|---|---|---|---|---|
| | 0,117 | 0,411 | 0,507 | 0,706 | 0,994 | 0,999 |

**3)** **α)** Για δύο φυσικούς αριθμούς κ και ν με κ ≤ ν, να αποδείξετε ότι:

$$\binom{\nu}{\kappa}=\binom{\nu}{\nu-\kappa}$$

**β)** Πώς ερμηνεύετε συνδυαστικά την παραπάνω ισότητα;

**γ)** Να αποδείξετε ότι $\binom{\nu}{\kappa}=\binom{\nu-1}{\kappa}+\binom{\nu-1}{\kappa-1}$.

**δ)** Μπορείτε να ερμηνεύσετε συνδυαστικά την ισότητα του (γ) χρησιμοποιώντας το ακόλουθο πρόβλημα;

> «Από ν όμοια σφαιρίδια που βρίσκονται σε ένα δοχείο σημαδεύουμε ένα με μαρκαδόρο. Με πόσους τρόπους μπορούμε να επιλέξουμε τυχαία κ από τα ν σφαιρίδια»;

# ΣΤΑΤΙΣΤΙΚΗ

*Εισαγωγή*

Πολλές φορές ακούμε από τα μέσα μαζικής ενημέρωσης ή διαβάζουμε στα διάφορα έντυπα για τα αποτελέσματα μιας έρευνας που βασίζεται σε στατιστικά δεδομένα. Αυτές οι πληροφορίες, που αναφέρονται σε συγκεκριμένη έρευνα, συνήθως δίνονται με τη μορφή πινάκων ή διαγραμμάτων. Εκτός από την παρουσίαση αυτών των πληροφοριών, πολλές φορές ακούμε ή διαβάζουμε μελέτες και αναλύσεις των παρατηρήσεων που αναφέρονται στην έρευνα.

Στατιστική είναι η εφαρμοσμένη επιστήμη που συγκεντρώνει και παρουσιάζει πληροφορίες, αλλά ταυτόχρονα μελετά και αναλύει τις παρατηρήσεις, που αναφέρονται στην οικονομία, στη μετεωρολογία, στην υγεία, στην κοινωνιολογία, στον αθλητισμό κτλ.

Η Στατιστική περιλαμβάνει τόσο τις μεθόδους συλλογής και επεξεργασίας στοιχείων όσο και τις μεθόδους ανάλυσης και μελέτης αυτών. Από τη μελέτη αυτή προκύπτουν οι σχέσεις που υπάρχουν στα διάφορα φαινόμενα και διατυπώνονται συμπεράσματα που είναι χρήσιμα για τη λήψη αποφάσεων.

Ως ορισμό της «Στατιστικής» θα δώσουμε τον συνηθέστερο και πλέον γνωστό ορισμό του R.A. Fisher (1890-1962), πατέρα της σύγχρονης Στατιστικής:

Στατιστική είναι ένα σύνολο αρχών και μεθοδολογιών για:

- τον σχεδιασμό της διαδικασίας συλλογής δεδομένων
- τη συνοπτική και αποτελεσματική παρουσίασή τους
- την ανάλυση και εξαγωγή αντίστοιχων συμπερασμάτων.

Ο κλάδος της Στατιστικής που ασχολείται με τον πρώτο στόχο λέγεται **σχεδιασμός πειραμάτων**, ενώ με τον δεύτερο ασχολείται η **περιγραφική στατιστική**, που αποτελεί και το αντικείμενο μελέτης μας στη συνέχεια. Τέλος, η **επαγωγική στατιστική ή στατιστική συμπερασματολογία** περιλαμβάνει τις μεθόδους με τις οποίες γίνεται η προσέγγιση των χαρακτηριστικών ενός μεγάλου συνόλου δεδομένων, με τη μελέτη των χαρακτηριστικών ενός μικρού υποσυνόλου των δεδομένων.

## ΕΝΟΤΗΤΑ 2.1 : ΠΛΗΘΥΣΜΟΣ - ΔΕΙΓΜΑ - ΜΕΤΑΒΛΗΤΕΣ

### Διερεύνηση

Δυο φοιτητές κοινωνιολογίας έχουν ως θέμα εργασίας να μελετήσουν κατά πόσον οι νέοι ηλικίας μεταξύ 12 και 20 ετών όλης της χώρας συζητούν στα μέσα κοινωνικής δικτύωσης θέματα που έχουν να κάνουν με την προσωπική τους ζωή και αν το γεγονός αυτό επιφέρει αρνητικές επιπτώσεις στους ίδιους. Για να βοηθηθούν στην έρευνά τους, συγκέντρωσαν από το διαδίκτυο τα αποτελέσματα παρόμοιων ερευνών που πραγματοποιήθηκαν σε άλλες χώρες.


Σύμφωνα με τα στοιχεία μιας έρευνας που διεξήχθη πρόσφατα (πηγή ΑΠΕ-ΜΠΕ), τα κοινωνικά δίκτυα δημιουργούν επιπρόσθετο στρες στα νέα παιδιά της Βρετανίας. Για τη διεξαγωγή της έρευνας ερωτήθηκαν 5.000 νέοι/ες, ηλικίας μεταξύ 12 και 20 ετών, από όλη τη χώρα. Εκ των ερωτηθέντων, το 40% δήλωσε ότι αισθάνεται άσχημα όταν κανείς από τους «φίλους» τους δεν κάνει «like» στις «selfies» τους, ενώ το 35% απάντησε ότι το επίπεδο της αυτοπεποίθησής τους συνδέεται άμεσα από τον αριθμό των «ακολούθων» που έχουν. Περισσότερες λεπτομέρειες για τα αποτελέσματα της συγκεκριμένης έρευνας φαίνονται στη διπλανή εικόνα.

***Θέματα για συζήτηση***

1) Ποιους/ες εκπροσωπούν οι 5.000 νέοι/ες της έρευνας;
2) Ποια μπορεί να ήταν τα ερωτήματα και ποια τα χαρακτηριστικά ως προς τα οποία εξετάστηκαν οι ερωτηθέντες;

Στη συνέχεια, οι δυο φοιτητές ετοίμασαν ένα αντίστοιχο ερωτηματολόγιο για τη δική τους έρευνα και το έστειλαν στους φίλους τους και τους συμφοιτητές τους για να το απαντήσουν.

***Ερωτήσεις***

Να γράψετε μερικές από τις ερωτήσεις που ενδεχομένως να συμπεριέλαβαν στο ερωτηματολόγιό τους οι δυο φοιτητές, οι οποίες να επιδέχονται άλλες ποσοτική και άλλες ποιοτική απάντηση.

1) Να σχολιάσετε τον τρόπο που εργάστηκαν οι φοιτητές για να συλλέξουν τα δεδομένα τους. Ήταν κατάλληλο το δείγμα που επέλεξαν ώστε να είναι αξιόπιστη η έρευνα;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

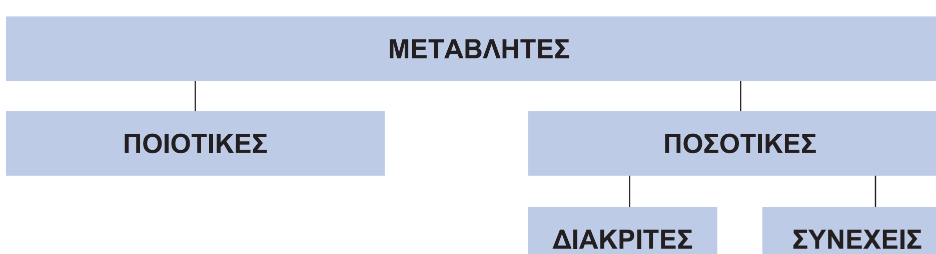
*Ορισμοί*

- Το σύνολο, στο οποίο επικεντρωνόμαστε σε μια έρευνα, ονομάζεται **πληθυσμός** της έρευνας. Τα στοιχεία του πληθυσμού συχνά αναφέρονται ως **άτομα** του πληθυσμού.
- Τις περισσότερες φορές είναι πρακτικά πολύ δύσκολο ή και αδύνατο να προσεγγίσουμε καθένα από τα άτομα του πληθυσμού, για τεχνικούς λόγους ή λόγω χρόνου, κόστους κτλ. Στις περιπτώσεις αυτές επιλέγουμε ένα μέρος του πληθυσμού που είναι αντιπροσωπευτικό του πληθυσμού και το εξετάζουμε. Το μέρος αυτό του πληθυσμού λέγεται **δείγμα**.
- Τα χαρακτηριστικά ως προς τα οποία εξετάζουμε τα άτομα ενός πληθυσμού λέγονται **μεταβλητές** και συμβολίζονται με ένα κεφαλαίο γράμμα X, Y, Z, ... κτλ.

Οι μεταβλητές διακρίνονται σε:

*Διάκριση μεταβλητών*

- **Ποιοτικές** αν αναφέρονται σε ένα ποιοτικό χαρακτηριστικό του πληθυσμού, όπως είναι: τα χρώματα των αυτοκινήτων, το φύλο των μαθητών κτλ.
- **Ποσοτικές** αν αναφέρονται σε ένα ποσοτικό χαρακτηριστικό του πληθυσμού, όπως είναι: ο ετήσιος αριθμός των τροχαίων ατυχημάτων, το ύψος των μαθητών κτλ.
    - **Διακριτές** αν παίρνουν μεμονωμένες τιμές, όπως είναι: ο αριθμός των παιδιών των οικογενειών, το νούμερο των γυναικείων παπουτσιών ανά μισό πόντο κτλ.
    - **Συνεχείς** αν μπορούν να πάρουν οποιαδήποτε τιμή ενός διαστήματος, όπως είναι: το βάρος των μαθητών, ο χρόνος που χρειάζονται οι μαθητές για να απαντήσουν σε ένα διαγώνισμα κτλ.




## Εφαρμογές

### Εφαρμογή 1

Μια ομάδα μαθητών θέλει να προβεί σε πρόβλεψη για το ποιος θα είναι ο νέος πρόεδρος του μαθητικού συμβουλίου του σχολείου. Για τον λόγο αυτό τα μέλη της ομάδας κατέγραψαν την πρόθεση ψήφου των μαθητών ενός συγκεκριμένου τμήματος του σχολείου.

1) Ποιος είναι ο πληθυσμός της έρευνας;

2) Ποιο είναι το δείγμα;
3) Ποια είναι η μεταβλητή της έρευνας και ποιο το είδος της;
4) Να συζητήσετε πλεονεκτήματα και μειονεκτήματα της μεθόδου που επέλεξαν οι μαθητές, για να συγκεντρώσουν τις πληροφορίες και να κάνουν την πρόβλεψή τους.

**Λύση**

1) Ο πληθυσμός της έρευνας είναι όλοι οι μαθητές του σχολείου.
2) Το δείγμα αποτελείται από τους μαθητές του τμήματος των οποίων καταγράφτηκε η ψήφος τους.
3) Μεταβλητή της έρευνας είναι: «ο υποψήφιος που είναι προτιμώμενος για πρόεδρος του μαθητικού συμβουλίου». Η μεταβλητή αυτή είναι ποιοτική.
4) Το πλεονέκτημα της μεθόδου που επέλεξαν οι μαθητές είναι η ευκολία και η συντομία στο να πάρουν τις απαραίτητες απαντήσεις από τα άτομα του δείγματος. Το μειονέκτημα είναι ότι το δείγμα δεν είναι και αντιπροσωπευτικό του πληθυσμού, αφού αποτελείται από τους μαθητές ενός συγκεκριμένου τμήματος. Έτσι τα αποτελέσματα της έρευνας δε θα είναι αξιόπιστα, διότι οι υπόλοιποι μαθητές μπορεί να έχουν διαφορετική γνώμη.

### Εφαρμογή 2

Η τροχαία θέλει να κάνει μια έρευνα για τα τροχαία ατυχήματα που σημειώθηκαν το 2017 στον αυτοκινητόδρομο Αθηνών - Θεσσαλονίκης. Θέλει να εξετάσει τις πιο κάτω μεταβλητές:
1) τον τύπο του οχήματος που έχει εμπλακεί στο δυστύχημα,
2) την ταχύτητα του οχήματος,
3) τον αριθμό των τραυματιών,

Να χαρακτηρίσετε το είδος καθεμιάς από τις μεταβλητές.

**Λύση**

Οι μεταβλητές χαρακτηρίζονται ως προς το είδος τους ως ακολούθως:
1) ποιοτική,
2) ποσοτική συνεχής,
3) ποσοτική διακριτή.

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Ο δήμαρχος μιας πόλης θέλει να διερευνήσει τα προβλήματα που αντιμετωπίζει η πόλη του, ώστε να δώσει έμφαση στην επίλυση αυτών. Για τον λόγο αυτό, ανέθεσε σε μια εταιρεία δημοσκοπήσεων μια έρευνα κατά την οποία 500 δημότες κλήθηκαν να δηλώσουν ποιο πρόβλημα της πόλης θεωρούν το πιο σημαντικό. Στις παρακάτω ερωτήσεις να επιλέξετε τη σωστή απάντηση.

1) Ο πληθυσμός της έρευνας είναι:
   1) Οι 500 δημότες που ρωτήθηκαν.
   2) Όλοι οι δημότες της πόλης.
   3) Όλοι οι Έλληνες πολίτες.
2) Το δείγμα αποτελούν:
   1) Οι 500 δημότες που ρωτήθηκαν.
   2) Όλοι οι δημότες της πόλης.
   3) Όλοι οι Έλληνες πολίτες.
3) Η μεταβλητή της έρευνας είναι:
   1) Τα προβλήματα της πόλης.
   2) Τα προβλήματα των δημοτών.
   3) Το σημαντικότερο πρόβλημα της πόλης.

**2)** Στις παρακάτω περιπτώσεις ποιες μπορεί να είναι οι μεταβλητές που μας ενδιαφέρουν; Να γίνει η διάκρισή τους σε ποιοτικές ή ποσοτικές και να αναφερθούν μερικές δυνατές τιμές τους:
1) Εξετάζουμε ένα δείγμα υπαλλήλων μιας εταιρείας.
2) Εξετάζουμε ένα δείγμα προϊόντων από μια παραγωγή.
3) Εξετάζουμε ένα δείγμα τηλεθεατών.
4) Εξετάζουμε τους καλαθοσφαιριστές μιας ομάδας σε έναν αγώνα.

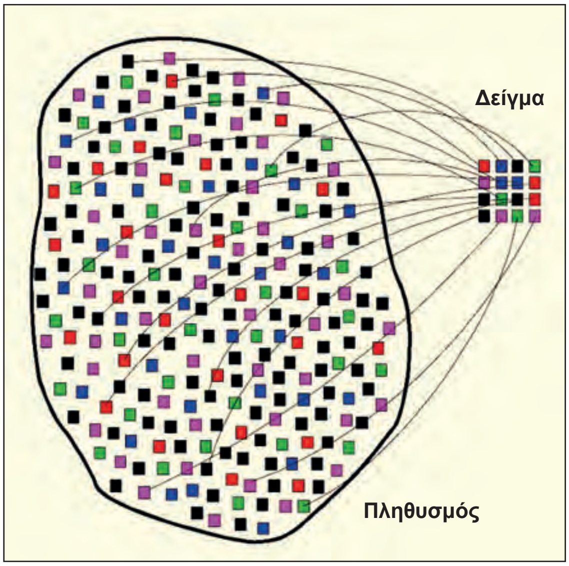
**3)** Για τις ανάγκες μιας έρευνας συγκεντρώσαμε στοιχεία από διερχόμενα οχήματα, σε κάποιο συγκεκριμένο σημείο της πόλης, κατά τη διάρκεια ενός εικοσιτετραώρου. Στον επόμενο πίνακα παρουσιάζονται μερικά από τα στοιχεία αυτά.

| **Είδος οχήματος** | **Χρώμα οχήματος** | **Ταχύτητα σε Km/h** | **Πλήθος επιβατών** |
|---|---|---|---|
| Φορτηγό | Κόκκινο | 26 | 2 |
| Αυτοκίνητο ΙΧ | Γκρι | 38 | 4 |
| Ποδήλατο | Πράσινο | 13 | 1 |
| Αυτοκίνητο ΙΧ | Κόκκινο | 52 | 2 |
| Λεωφορείο | Λευκό | 34 | 25 |
| Αυτοκίνητο ΙΧ | Λευκό | 45 | 1 |
| Μοτοσικλέτα | Μαύρο | 62 | 2 |

Ποιες είναι οι μεταβλητές της έρευνας και ποιο το είδος τους;

**4)** Σε μια εκπομπή δημόσιας συζήτησης, συγκεκριμένου τηλεοπτικού καναλιού, το κοινό καλείται να ψηφίσει αν συμφωνεί με την Α άποψη ή τη Β άποψη. Να αναφέρετε δύο λόγους για τους οποίους τα αποτελέσματα της ψηφοφορίας δεν μπορεί να γενικευτούν σε ολόκληρο τον πληθυσμό της χώρας.

## Πρόσθετο υλικό




Η διαδικασία επιλογής ενός δείγματος από τον πληθυσμό ονομάζεται **δειγματοληψία**. Όπως είπαμε, προκειμένου να γενικεύσουμε τα συμπεράσματα που θα εξαγάγουμε από το δείγμα σε ολόκληρο τον πληθυσμό, έχει ιδιαίτερη σημασία το δείγμα να είναι αντιπροσωπευτικό του πληθυσμού. Αυτό σημαίνει ότι τα χαρακτηριστικά που θεωρούμε ότι επηρεάζουν τις μεταβλητές που μας ενδιαφέρουν κατανέμονται με τον ίδιο τρόπο στο δείγμα και στον πληθυσμό. Στο παράδειγμα της Εφαρμογής 1, η σωστή πρακτική θα ήταν να συμπεριλάβουμε στο δείγμα μας μαθητές από κάθε τμήμα. Θα ήταν επίσης σωστό να συμπεριλάβουμε στο δείγμα μας αγόρια και κορίτσια με την ίδια περίπου αναλογία που βρίσκονται στον σχολικό πληθυσμό.

Ένας απλός τρόπος δειγματοληψίας που επιτυγχάνει στατιστικά την αντιπροσωπευτικότητα είναι η απλή τυχαία δειγματοληψία, στην οποία επιλέγουμε τυχαία τα μέλη του πληθυσμού που θα συμπεριληφθούν στο δείγμα. Για παράδειγμα, αν από έναν πληθυσμό 4.000.000 ανδρών και 4.000.000 γυναικών επιλέξουμε τυχαία 1.000 άτομα, τότε με πιθανότητα περίπου 99% το πλήθος των γυναικών που επιλέξαμε θα είναι μεταξύ 460 και 540.

Όσο μεγαλύτερο είναι το μέγεθος του δείγματος, τόσο πιο βέβαιοι μπορούμε να είμαστε ότι τα αποτελέσματα που θα πάρουμε από το δείγμα μας γενικεύονται στον πληθυσμό.

### Ιστορικά στοιχεία

Στις προεδρικές εκλογές στις ΗΠΑ το 1936 υπήρξε πλήρης αποτυχία της πρόγνωσης των εκλογικών αποτελεσμάτων από το περιοδικό «Literary Digest». Η έρευνα, που διενήργησε το περιοδικό, στηρίχτηκε σε αρχικό δείγμα 10.000.000 ατόμων του πληθυσμού από τηλεφωνικούς καταλόγους, τον κατάλογο των συνδρομητών του περιοδικού και καταλόγους ιδιοκτητών αυτοκινήτων. Στο ερωτηματολόγιο απάντησαν περίπου 2.400.000 άτομα και από την ανάλυση που πραγματοποιήθηκε φάνηκε ότι το εκλογικό σώμα θα έδινε μεγάλη πλειοψηφία στους Ρεπουμπλικάνους με υποψήφιο πρόεδρο τον Landon. Η μέρα τον εκλογών όμως επιφύλαξε μια δυσάρεστη έκπληξη στους στατιστικούς αναλυτές της δημοσκόπησης. Ο Δημοκρατικός υποψήφιος Roosevelt εκλέχτηκε με την ιστορική πλειοψηφία 60,8%, σε αντίθεση με το 42,9% που είχε προβλέψει η δημοσκόπηση του «Literary Digest».

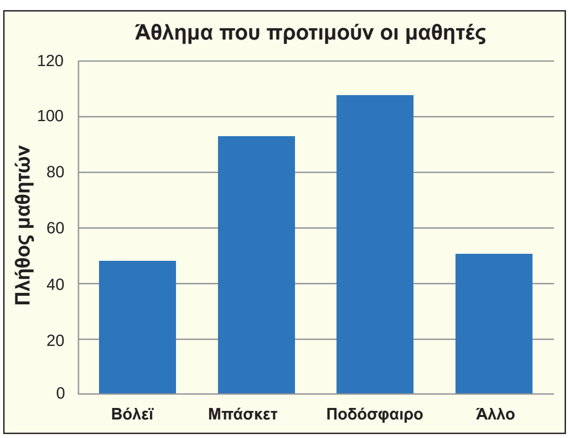
Γιατί νομίζετε η έρευνα απέτυχε και μάλιστα σε τόσο μεγάλο βαθμό;

# ΕΝΟΤΗΤΑ 2.2 : ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

## Διερεύνηση

Μια εταιρεία αθλητικών ειδών, πριν επενδύσει σε αθλητικά είδη που προτιμούν οι μαθητές, αποφάσισε να κάνει μια έρευνα. Για τον λόγο αυτό επέλεξε, με τυχαίο τρόπο, δείγμα τριακοσίων μαθητών απ' όλη την Ελλάδα. Ο υπεύθυνος που έκανε την έρευνα, μετά την επεξεργασία των στοιχείων που συγκέντρωσε, παρουσίασε στον διευθυντή της εταιρείας τον παρακάτω πίνακα και τα δυο διαγράμματα.

| **Άθλημα** | **Πλήθος μαθητών** | **Ποσοστό** |
|---|---|---|
| Βόλεϊ | 48 | 16% |
| Μπάσκετ | 93 | 31% |
| Ποδόσφαιρο | 108 | 36% |
| Άλλο | 51 | 17% |
| **Σύνολο** | **300** | **100%** |





1) Ποια είναι η μεταβλητή της έρευνας και ποιο το είδος της;
2) Πώς προκύπτουν τα αντίστοιχα ποσοστά για κάθε άθλημα;
3) Ποιο είναι το ύψος της κάθε μπάρας στο ραβδόγραμμα;
4) Ποια είναι η γωνία του κάθε κυκλικού τομέα στο κυκλικό διάγραμμα;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

*Ορισμοί*

- Έστω X μια μεταβλητή με τιμές $x_1$, $x_2$, $x_3$, ..., $x_\kappa$. **Συχνότητα $\nu_i$** μιας τιμής $x_i$ λέγεται ο φυσικός αριθμός που εκφράζει πόσες φορές εμφανίζεται η τιμή αυτή στο σύνολο των παρατηρήσεων του δείγματος. Είναι φανερό ότι το άθροισμα όλων των συχνοτήτων μας δίνει το μέγεθος ν του δείγματος. Δηλαδή:

$$\boxed{\nu_1+\nu_2+\nu_3+\ldots\nu_\kappa=\nu}$$

- **Η σχετική συχνότητα $f_i$** μιας τιμής $x_i$ ορίζεται ως ο λόγος της αντίστοιχης συχνότητας $\nu_i$ προς το μέγεθος ν του δείγματος. Δηλαδή:

$$\boxed{f_i = \frac{\nu_i}{\nu} \text{ για } i = 1, 2, 3, \ldots, \kappa}$$

- Σύμφωνα με τους παραπάνω ορισμούς, ο πίνακας του παραδείγματος που αναφέραμε με μεταβλητή «το άθλημα που προτιμούν οι μαθητές» μετατρέπεται, όπως φαίνεται στη συνέχεια και λέγεται πίνακας **κατανομής συχνοτήτων και σχετικών συχνοτήτων**.

*Η σχετική συχνότητα μπορεί να εκφραστεί και ως ποσοστό $f_i$%.*

| **Άθλημα** $x_i$ | **Συχνότητα** $\nu_i$ | **Σχετική συχνότητα** $f_i$ | **Σχετική συχνότητα %** $f_i$% |
|---|---|---|---|
| Βόλεϊ | 48 | 0,16 | 16 |
| Μπάσκετ | 93 | 0,31 | 31 |
| Ποδόσφαιρο | 108 | 0,36 | 36 |
| Άλλο | 51 | 0,17 | 17 |
| **Σύνολο** | **300** | **1,00** | **100** |

***Γραφικές παραστάσεις***

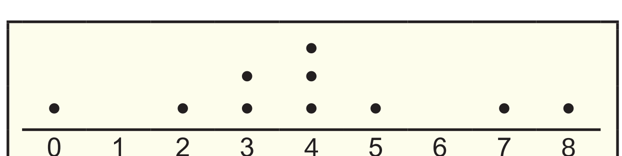
- Οι πλέον συνηθισμένοι τρόποι γραφικής παρουσίασης ποιοτικών αλλά και ποσοτικών διακριτών δεδομένων είναι το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα.

  - **Το ραβδόγραμμα συχνοτήτων** αποτελείται από ορθογώνιες στήλες, μια για κάθε τιμή της μεταβλητής, όπου το ύψος της κάθε στήλης είναι ίσο με την αντίστοιχη συχνότητα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **ραβδόγραμμα σχετικών συχνοτήτων**.

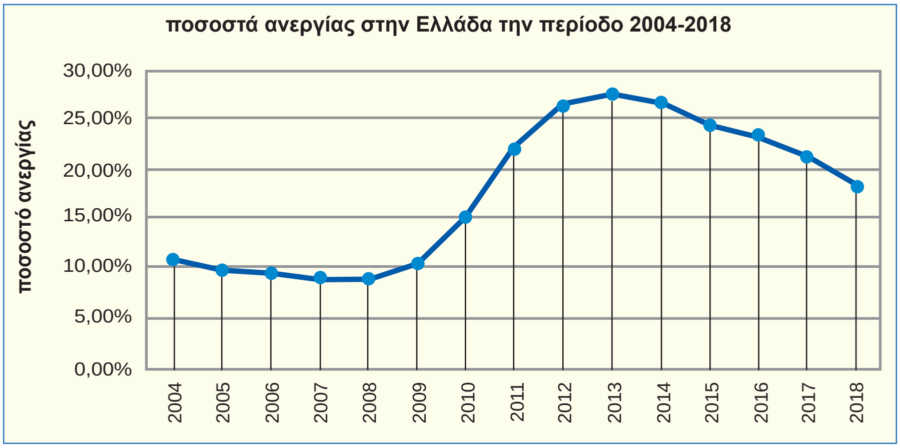
  - **Το κυκλικό διάγραμμα** χρησιμοποιείται επίσης για τη γραφική παράσταση δεδομένων, κυρίως όταν αυτά παίρνουν λίγες τιμές. Η γωνία του κάθε κυκλικού τομέα, είναι ανάλογη της αντίστοιχης σχετικής συχνότητας. Δηλαδή:

$$\boxed{\alpha_i = 360° \cdot f_i \text{ για } i = 1, 2, 3, \ldots, \kappa}$$

- Όταν έχουμε λίγες παρατηρήσεις, τότε η κατανομή τους μπορεί να περιγραφεί με το **σημειόγραμμα** στο οποίο οι τιμές παριστάνονται με σημεία υπεράνω ενός άξονα.



- Το **χρονόγραμμα** χρησιμοποιείται για την γραφική απεικόνιση της εξέλιξης σε σχέση με το χρόνο ενός μεγέθους, συνήθως οικονομικού ή δημογραφικού.



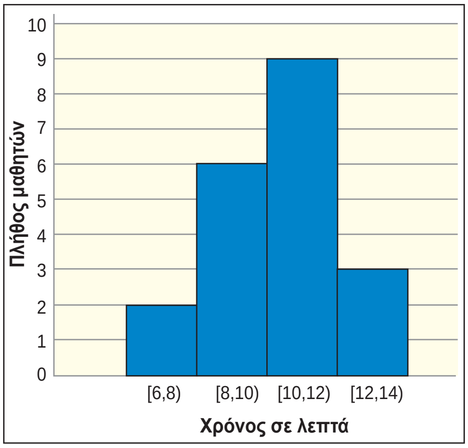

***Δεδομένα σε κλάσεις***

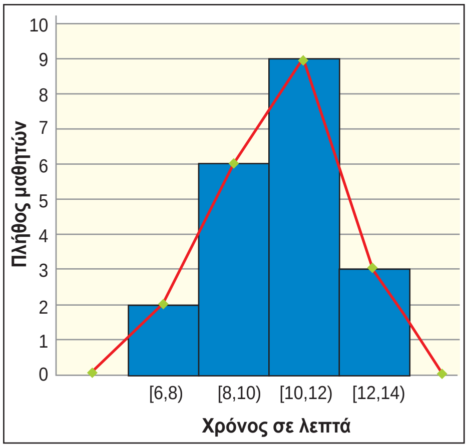
- Στη περίπτωση που έχουμε ποσοτικά συνεχή δεδομένα με πολλές διαφορετικές τιμές, τότε, για την καλύτερη παρουσίασή τους, γίνεται **ομαδοποίηση** αυτών σε κλάσεις, συνήθως ίσου πλάτους. Σχετικός είναι ο παρακάτω πίνακας με τους χρόνους που χρειάστηκαν οι 20 μαθητές ενός τμήματος για να απαντήσουν σε ένα πρόβλημα.

| **Χρόνος σε λεπτά των μαθητών** | **Συχνότητα $\nu_i$** | **Σχετική συχνότητα $f_i$** | **Σχετική συχνότητα % $f_i\%$** |
|---|---|---|---|
| [6,8) | 2 | 0,10 | 10 |
| [8,10) | 6 | 0,30 | 30 |
| [10,12) | 9 | 0,45 | 45 |
| [12,14) | 3 | 0,15 | 15 |
| **Σύνολο** | **20** | **1,00** | **100** |

- Η γραφική παρουσίαση ομαδοποιημένων στατιστικών δεδομένων γίνεται με το **ιστόγραμμα συχνοτήτων**. Στο παρακάτω σχήμα φαίνεται η γραφική παρουσίαση των δεδομένων του προηγουμένου πίνακα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **ιστόγραμμα σχετικών συχνοτήτων**.

- Αν θεωρήσουμε δυο επιπλέον κλάσεις ίσου πλάτους, μια στην αρχή και μια στο τέλος, με συχνότητα 0 και ενώσουμε τα μέσα των άνω βάσεων, τότε προκύπτει το λεγόμενο **πολύγωνο συχνοτήτων**. Σχετικό είναι το παρακάτω σχήμα. Αν αντί για τις συχνότητες έχουμε τις σχετικές συχνότητες, τότε λέγεται **πολύγωνο σχετικών συχνοτήτων**.




## Εφαρμογές

### Εφαρμογή 1

Σε μια δημοσκόπηση που έγινε για τις δημοτικές εκλογές, 280 άτομα απάντησαν ότι προτιμούν τον υποψήφιο «Α», 320 άτομα τον υποψήφιο «Β» και 200 άτομα τον υποψήφιο «Γ».

1) Ποιο ήταν το μέγεθος του δείγματος;
2) Να κάνετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
3) Να παραστήσετε τα δεδομένα με ραβδόγραμμα αλλά και με κυκλικό διάγραμμα.

**Λύση**

1) Το πλήθος αυτών που απάντησαν είναι:

$$\nu=280+320+200=800$$

Επομένως το μέγεθος του δείγματος είναι 800 άτομα.

2) Οι συχνότητες είναι: $\nu_1 = 280$, $\nu_2 = 320$ και $\nu_3 = 200$. Οι σχετικές συχνότητες είναι:

$$f_1 = \frac{\nu_1}{\nu} = \frac{280}{800} = 0{,}35,\quad f_2 = \frac{\nu_2}{\nu} = \frac{320}{800} = 0{,}40 \text{ και } f_3 = \frac{\nu_3}{\nu} = \frac{200}{800} = 0{,}25$$

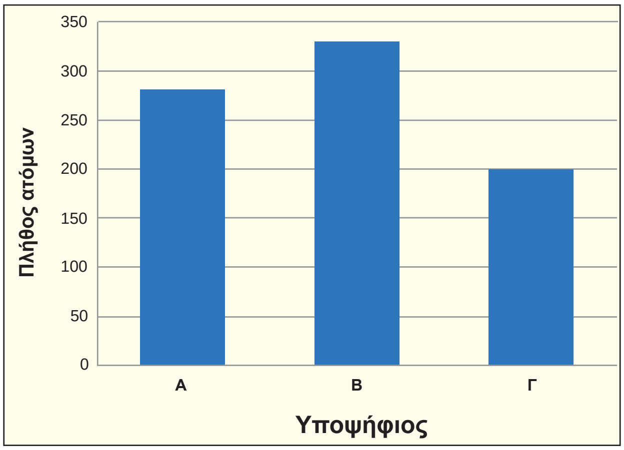
Ο πίνακας κατανομής συχνοτήτων και σχετικών συχνοτήτων είναι:

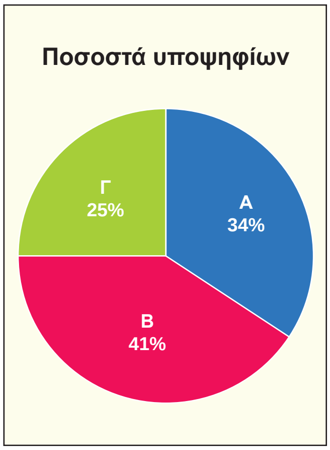
| Υποψήφιος δήμαρχος | Συχνότητα $\nu_i$ | Σχετική συχνότητα $f_i$ | Σχετική συχνότητα % $f_i\%$ |
|---|---|---|---|
| Υποψήφιος «Α» | 280 | 0,35 | 35 |
| Υποψήφιος «Β» | 320 | 0,40 | 40 |
| Υποψήφιος «Γ» | 200 | 0,25 | 25 |
| **Σύνολο** | **800** | **1,00** | **100** |

3) Για τις γωνίες των κυκλικών τομέων στο κυκλικό διάγραμμα έχουμε:

$$\alpha_1 = 360° \cdot f_1 = 360° \cdot 0{,}35 = 126°$$
$$\alpha_2 = 360° \cdot f_2 = 360° \cdot 0{,}40 = 144°$$
$$\alpha_3 = 360° \cdot f_3 = 360° \cdot 0{,}25 = 90°$$

Παρακάτω φαίνονται το ραβδόγραμμα συχνοτήτων και το κυκλικό διάγραμμα.




## Εφαρμογή 2

Παρακάτω δίνονται οι χρόνοι, στρογγυλοποιημένοι στο δέκατο του δευτερολέπτου, που απαιτήθηκαν για να τρέξουν 50 αθλητές έναν αγώνα δρόμου 400 m.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 52,1 | 55,3 | 50 | 56,4 | 59,1 | 54,2 | 56,7 | 54,4 | 57,1 | 53,7 |
| 55,2 | 55,1 | 58 | 59,2 | 56 | 55,5 | 52,5 | 56,5 | 58,5 | 55 |
| 55,2 | 57,3 | 54,3 | 53,5 | 57,9 | 53 | 55,4 | 55,6 | 52,4 | 54,5 |
| 56,4 | 59,1 | 54,2 | 56,7 | 55,3 | 52,4 | 56,4 | 54,1 | 54,3 | 56,7 |
| 54,3 | 51,5 | 57 | 53,2 | 54,9 | 55,6 | 52 | 55,3 | 55,1 | 54,7 |

1) Ποιο είναι το είδος της μεταβλητής, ποιος ο μικρότερος και ποιος ο μεγαλύτερος χρόνος;
2) Τι έχετε να παρατηρήσετε για το πλήθος των τιμών των παρατηρήσεων σε σχέση με το πλήθος των παρατηρήσεων;
3) Ξεκινώντας από το μικρότερο χρόνο και με βήμα 2 sec, ποιες κλάσεις της μορφής [α,β) δημιουργούνται στις οποίες περιέχονται όλες οι παρατηρήσεις;
4) Να παρουσιαστούν οι παρατηρήσεις, ομαδοποιημένες στις παραπάνω κλάσεις ίσου πλάτους, σε έναν πίνακα συχνοτήτων και σχετικών συχνοτήτων.
5) Να παραστήσετε τα δεδομένα με ιστόγραμμα συχνοτήτων και με πολύγωνο συχνοτήτων.

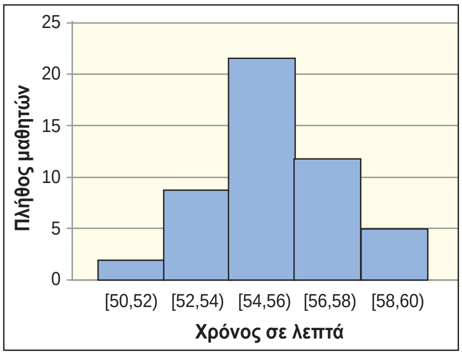
**Λύση**

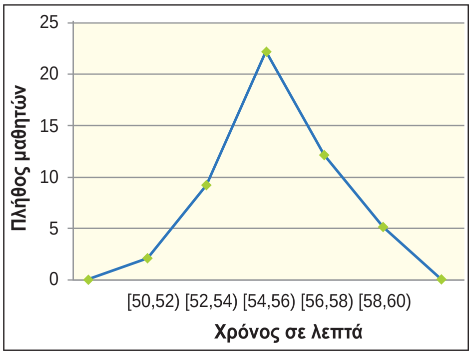
1) Η μεταβλητή είναι ποσοτική συνεχής. Ο μικρότερος χρόνος είναι 50 sec και ο μεγαλύτερος 59,2 sec.

2) Το πλήθος των τιμών των παρατηρήσεων είναι 37 σε σύνολο 50 παρατηρήσεων. Για να παρουσιαστούν τα δεδομένα, ως έχουν, σε ένα πίνακα συχνοτήτων θα χρειαστούμε 37 γραμμές. Είναι φανερό ότι ένας τέτοιος πίνακας είναι δύσχρηστος και οι πληροφορίες δεν παρουσιάζονται συνοπτικά.
3) Οι κλάσεις που δημιουργούνται είναι: [50, 52), [52, 54), [54, 56), [56, 58) και [58, 60)
4) Μετά τη διαλογή ο πίνακας συχνοτήτων και σχετικών συχνοτήτων είναι ο παρακάτω.

| **Κλάσεις με χρόνους σε sec** | **Συχνότητα $v_i$** | **Σχετική συχνότητα $f_i$** | **Σχετική συχνότητα % $f_i\%$** |
|---|---|---|---|
| [50, 52) | 2 | 0,04 | 4 |
| [52, 54) | 9 | 0,18 | 18 |
| [54, 56) | 22 | 0,44 | 44 |
| [56, 58) | 12 | 0,24 | 24 |
| [58, 60) | 5 | 0,10 | 10 |
| **Σύνολο** | **50** | **1,00** | **100** |

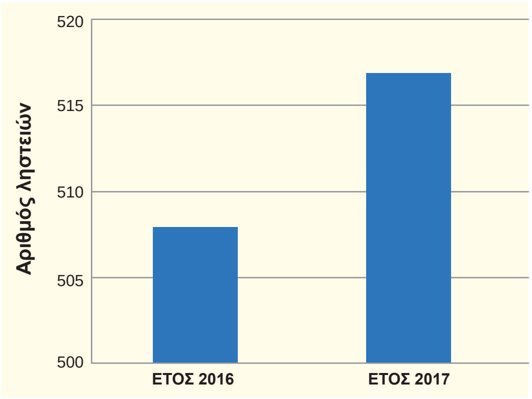
5) Παρακάτω φαίνονται το ιστόγραμμα και το πολύγωνο συχνοτήτων.




## Εφαρμογή 3

Σε ένα τηλεοπτικό κανάλι, ένας δημοσιογράφος σχολίασε τη διπλανή γραφική παράσταση ως εξής: *«Η γραφική παράσταση δείχνει ότι σημειώθηκε τεράστια αύξηση του αριθμού των ληστειών το έτος 2017 σε σχέση με το έτος 2016».*



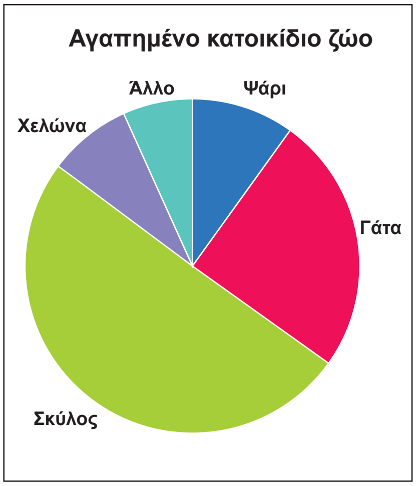

Νομίζετε ότι ο δημοσιογράφος του καναλιού αυτού ερμήνευσε σωστά την γραφική παράσταση; Να γράψετε ένα επιχείρημα που να τεκμηριώνει την απάντησή σας.

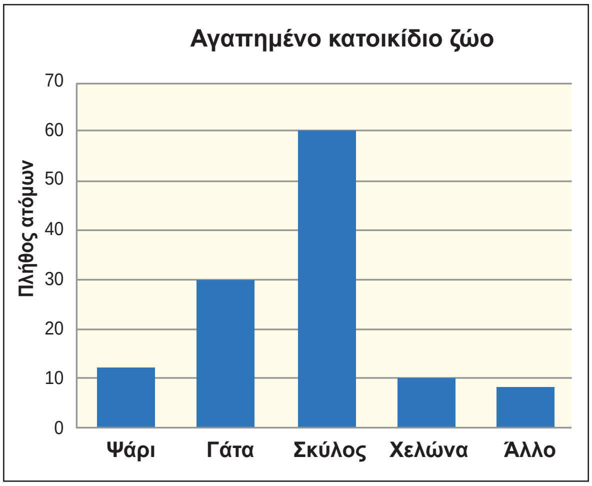
**Λύση**

Για τον λόγο ότι η δεύτερη στήλη φαίνεται περίπου διπλάσια στο ύψος από την πρώτη στήλη, δόθηκε η λανθασμένη ερμηνεία. Όμως αυτό που πραγματικά φαίνεται είναι το άνω τμήμα των στηλών, αφού ο άξονας των τεταγμένων αρχίζει από το 500. Έτσι, το πραγματικό ύψος της πρώτης στήλης είναι 508, ενώ της δεύτερης είναι 517. Επομένως ο αριθμός των ληστειών αυξήθηκε μόνο κατά 9 σε σύνολο 500 και πλέον ληστειών. Το γεγονός αυτό δε δικαιολογεί την παραπάνω ερμηνεία.

## Εφαρμογή 4

Ρωτήθηκαν 120 άτομα για το αγαπημένο τους κατοικίδιο ζώο και τα αποτελέσματα παρουσιάζονται με ένα κυκλικό διάγραμμα και ένα ραβδόγραμμα.




1) Να κάνετε μια ερώτηση που μπορεί να απαντηθεί πιο εύκολα από το ραβδόγραμμα.

2) Να κάνετε μια ερώτηση που μπορεί να απαντηθεί πιο εύκολα από το κυκλικό διάγραμμα.

**Λύση**

1) Για το ραβδόγραμμα μια ερώτηση με εύκολη απάντηση μπορεί να είναι: «Πόσες φορές περισσότερα είναι τα άτομα που το αγαπημένο τους κατοικίδιο ζώο είναι ο σκύλος σε σχέση με αυτά που είναι η χελώνα»;

2) Για το κυκλικό διάγραμμα μια ερώτηση με εύκολη απάντηση μπορεί να είναι: «Ποιο είναι το ποσοστό των ατόμων που το αγαπημένο τους κατοικίδιο ζώο είναι ο σκύλος»;

## Ασκήσεις – Προβλήματα – Δραστηριότητες

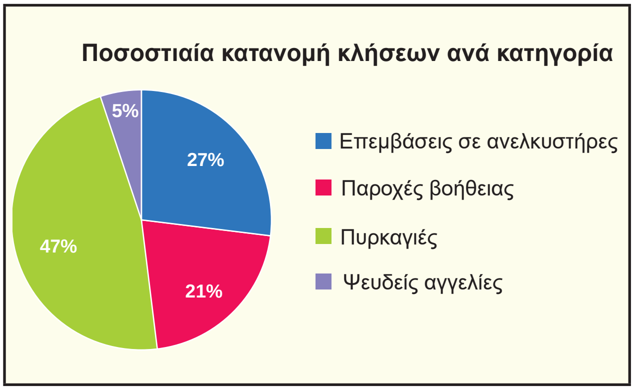
**1)** Το επάγγελμα του πατέρα 20 μαθητών καταγράφηκε στον διπλανό πίνακα. Να κάνετε πίνακα σχετικών συχνοτήτων και να κατασκευάσετε το αντίστοιχο ραβδόγραμμα συχνοτήτων καθώς και το κυκλικό διάγραμμα.

| Επάγγελμα πατέρα | Αριθμός ατόμων |
|---|---|
| Ιδιωτικός υπάλληλος | 6 |
| Δημόσιος υπάλληλος | 7 |
| Αυτοαπασχολούμενος | 5 |
| Άλλο | 2 |

**2)** Στον διπλανό πίνακα δίνονται τα καθαρά κέρδη μιας εταιρείας, ανά έτος, από το 2014 έως και το 2017. Να κάνετε χρονόγραμμα όπου να φαίνεται η εξέλιξη των κερδών σε σχέση με τον χρόνο.

| Έτος | Κέρδη σε ευρώ |
|---|---|
| 2014 | 180.000 |
| 2015 | 270.000 |
| 2016 | 230.000 |
| 2017 | 210.000 |

**3)** Η στατιστική υπηρεσία της Πυροσβεστικής μας έδωσε το διπλανό κυκλικό διάγραμμα, που παρουσιάζει τα ποσοστά των κλήσεων ανά κατηγορία. Αν το σύνολο των κλήσεων είναι 60.400, να γίνει πίνακας συχνοτήτων.




**4)** Δίνεται η ποσοστιαία σύνθεση (%) του προσωπικού μιας επιχείρησης, με αριθμό υπαλλήλων 80 άτομα ως προς το μορφωτικό τους επίπεδο.

1) Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.

2) Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με κυκλικό διάγραμμα.

| Τίτλος Σπουδών | Ποσοστό (%) |
|---|---|
| Μεταπτυχιακό Δίπλωμα | 20 |
| Πτυχίο Τριτοβάθμιας Εκπαίδευσης | 50 |
| Απολυτήριο Λυκείου | 30 |

**5)** Οι βαθμοί στην Ιστορία 25 μαθητών, ενός τμήματος της Β΄ τάξης ΓΕΛ, είναι:

1) Να κατασκευάσετε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων.
2) Να απεικονίσετε τα δεδομένα με ραβδόγραμμα συχνοτήτων και με σημειόγραμμα.

| | | | | |
|---|---|---|---|---|
| 16 | 15 | 17 | 16 | 17 |
| 18 | 17 | 16 | 17 | 18 |
| 16 | 19 | 17 | 15 | 16 |
| 17 | 16 | 15 | 17 | 18 |
| 17 | 14 | 17 | 16 | 19 |

**6)** Οι πιο κάτω αριθμοί παρουσιάζουν τις ενδείξεις της άνω έδρας ενός ζαριού το οποίο ρίξαμε 30 φορές.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 5 | 6 | 1 | 2 | 5 | 4 | 3 | 2 | 5 |
| 1 | 3 | 5 | 4 | 1 | 3 | 2 | 6 | 5 | 4 |
| 1 | 2 | 6 | 2 | 4 | 3 | 1 | 6 | 4 | 5 |

1) Να κατασκευάσετε πίνακα συχνοτήτων.
2) Να κατασκευάσετε ραβδόγραμμα συχνοτήτων.

**7)** Στον πιο κάτω πίνακα δίνεται η συγκέντρωση (mgr/cm$^3$) ενός ρύπου στον αέρα 40 πόλεων της χώρας.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 16 | 24 | 36 | 47 | 23 | 22 | 43 | 27 | 49 | 48 |
| 12 | 32 | 17 | 38 | 42 | 27 | 31 | 50 | 38 | 21 |
| 36 | 19 | 28 | 31 | 28 | 25 | 45 | 12 | 57 | 51 |
| 22 | 23 | 24 | 25 | 24 | 37 | 43 | 25 | 39 | 51 |

1) Να ομαδοποιήσετε τις παρατηρήσεις στις κλάσεις: [10,20), [20,30), [30,40), [40,50) και [50,60).
2) Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
3) Να κατασκευάσετε το ιστόγραμμα και το πολύγωνο συχνοτήτων.

**8)** Οι 50 εργάτες ενός εργοστασίου έχουν τις παρακάτω ηλικίες:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 21 | 43 | 50 | 25 | 55 | 30 | 28 | 40 | 31 | 51 |
| 18 | 47 | 52 | 34 | 47 | 32 | 27 | 41 | 35 | 54 |
| 30 | 48 | 36 | 43 | 38 | 33 | 27 | 39 | 41 | 43 |
| 32 | 22 | 46 | 52 | 29 | 32 | 34 | 34 | 42 | 36 |
| 35 | 28 | 57 | 56 | 20 | 38 | 27 | 27 | 40 | 35 |

1) Να ομαδοποιήσετε τις ηλικίες στις κλάσεις: [18,28), [28,38), [38,48) και [48,58).
2) Να κατασκευάσετε πίνακα συχνοτήτων και σχετικών συχνοτήτων.
3) Να κατασκευάσετε το ιστόγραμμα και το πολύγωνο συχνοτήτων.

## Πρόσθετο υλικό

Η εκμάθηση των βασικών εννοιών της στατιστικής γίνεται πιο ενδιαφέρουσα, όταν μπορούμε να εργαστούμε με πληροφορίες που συλλέγονται από εμάς και αποθηκεύονται σε έναν υπολογιστή. Η απλούστερη συλλογή δεδομένων μπορεί να αφορά πληροφορίες για τους συμμαθητές σας. Η έρευνα που ακολουθεί μπορεί να σας δώσει αρκετά δεδομένα, με τα οποία μπορείτε να απαντήσετε σε μερικές ενδιαφέρουσες ερωτήσεις. Αποθηκεύστε τα δεδομένα που θα συλλέξετε.

1) Φύλο

2) Ηλικία

3) Ύψος σε cm

4) Ύψος του πατέρα σας σε cm

5) Ύψος της μητέρας σας σε cm

6) Το πέμπτο ψηφίο του αριθμού της ταυτότητάς σας

7) Το τελευταίο ψηφίο του αριθμού ενός αυτοκινήτου της οικογένειας σας

8) Χρώμα μαλλιών

9) Χρώμα ματιών

Μπορείτε να προσθέσετε μια ή περισσότερες ερωτήσεις που σας ενδιαφέρουν. Μοιράστε ένα έντυπο ερωτηματολόγιου, όπου κάθε μαθητής θα γράφει τις απαντήσεις του. Επεξεργαστείτε τα δεδομένα που συγκεντρώσατε με ένα λογιστικό φύλλο.

1) Προσδιορίστε το είδος κάθε μεταβλητής

2) Κατασκευάστε τον πίνακα συχνοτήτων και σχετικών συχνοτήτων για κάθε μεταβλητή ξεχωριστά.

3) Κατασκευάστε τα αντίστοιχα διαγράμματα ανάλογα με το είδος της μεταβλητής.

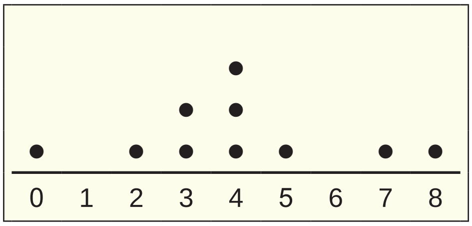
**<u>ΥΠΟΔΕΙΓΜΑ</u>**

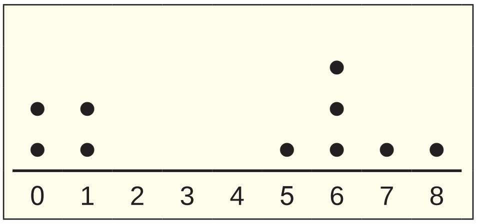
| α/α | Φύλο | Ηλικία | Ύψος | Ύψος πατέρα | Ύψος μητέρας | Τρίτο ψηφίο του αριθ. της ταυτότητας | Τελευταίο ψηφίο του αριθ. του αυτοκινήτου | Χρώμα μαλλιών | Χρώμα ματιών |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Θ | 17 | 167 | 170 | 160 | 8 | 1 | Μ | Μ |
| 2 | Θ | 18 | 165 | 172 | 163 | 4 | 7 | Κ | Μ |
| 3 | Α | 18 | 184 | 185 | 171 | 5 | 5 | Κ | Κ |
| 4 | Θ | 17 | 175 | 179 | 168 | 3 | 8 | Ξ | Μ |
| 5 | Α | 19 | 171 | 175 | 165 | 9 | 4 | Κ | Κ |
| 6 | Α | 18 | 169 | 170 | 167 | 6 | 7 | Μ | Κ |

# ΕΝΟΤΗΤΑ 2.3: ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ, ΘΗΚΟΓΡΑΜΜΑ, ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ

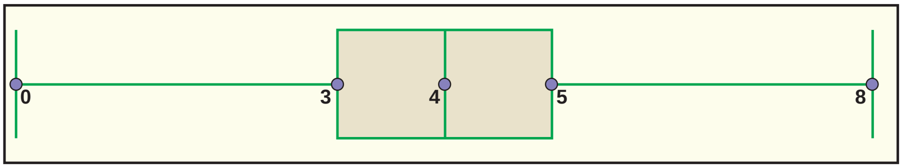
## Διερεύνηση

1) Στα επόμενα σημειογράμματα φαίνονται οι πόντοι που σημείωσαν οι 10 καλαθοσφαιριστές κάθε μιας από τις δυο ομάδες μπάσκετ, στο ημίχρονο ενός αγώνα.




   1) Να υπολογίσετε την μέση τιμή των πόντων που σημείωσαν οι 10 καλαθοσφαιριστές κάθε ομάδας στο ημίχρονο.
   2) Αφού γράψετε τις 10 παρατηρήσεις κάθε ομάδας σε αύξουσα σειρά, να βρείτε για κάθε ομάδα:
      1) Τη μικρότερη παρατήρηση.
      2) Τη μεσαία παρατήρηση από τις πέντε πρώτες, δηλαδή το πρώτο τεταρτημόριο.
      3) Το ημιάθροισμα των δυο μεσαίων παρατηρήσεων, δηλαδή το δεύτερο τεταρτημόριο ή αλλιώς τη διάμεσο.
      4) Τη μεσαία παρατήρηση από τις πέντε τελευταίες, δηλαδή το τρίτο τεταρτημόριο.
      5) Τη μεγαλύτερη παρατήρηση
   3) Με βάση τα παραπάνω, το θηκόγραμμα για την πρώτη ομάδα είναι αυτό που βλέπετε στο επόμενο σχήμα. Να σχεδιάσετε το θηκόγραμμα για τη δεύτερη ομάδα.



4) Ποια είναι η διαφορά της μικρότερης από τη μεγαλύτερη παρατήρηση, δηλαδή το εύρος των παρατηρήσεων, για κάθε ομάδα; Είναι το εύρος ένα αξιόπιστο μέτρο μεταβλητότητας;

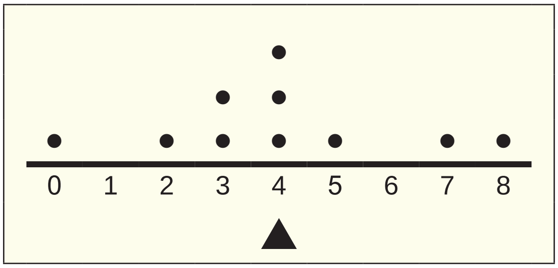
5) Ποια είναι η διαφορά του πρώτου από το τρίτο τεταρτημόριο, δηλαδή το ενδοτεταρτημοριακό εύρος, για κάθε ομάδα;

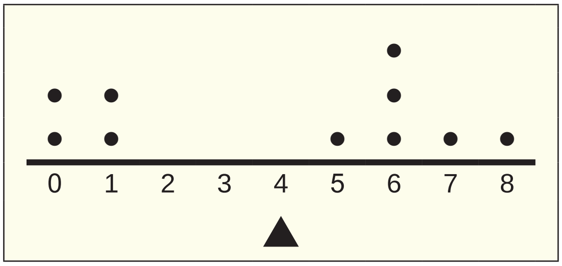
6) Ποια τιμή εμφανίζεται τις περισσότερες φορές, δηλαδή ποια είναι η επικρατούσα τιμή για κάθε ομάδα;

7) Πώς θα υπολογίσουμε τη μέση τιμή των τετραγώνων των διαφορών της μέσης τιμής των παρατηρήσεων από τις παρατηρήσεις, δηλαδή τη διακύμανση των παρατηρήσεων; Περιγράψτε τη διαδικασία και τις αριθμητικές πράξεις που απαιτούνται για κάθε ομάδα ξεχωριστά.

8) Να υπολογίσετε την τετραγωνική ρίζα της διακύμανσης, δηλαδή την τυπική απόκλιση για κάθε ομάδα.

2) Στα επόμενα δύο σχήματα οι κουκίδες συμβολίζουν δυο ομάδες, από 10 βαρίδια ίδιου βάρους η κάθε μια, τοποθετημένα πάνω σε δυο αβαρείς ράβδους αντίστοιχα, αριθμημένες έτσι ώστε οι αποστάσεις των διαδοχικών αριθμών να είναι ίσες.




1) Να εξηγήσετε τον λόγο για τον οποίο, αν τοποθετήσουμε μια τριγωνική σφήνα στη θέση 4 σε κάθε ράβδο, τότε αυτές ισορροπούν. Ποια είναι η φυσική ερμηνεία της μέσης τιμής των βαρών σε σχέση με το σημείο ισορροπίας αυτών;

2) Είναι το εύρος, στην παραπάνω περίπτωση, ένα αξιόπιστο μέτρο μεταβλητότητας των βαρών των δυο ομάδων;

3) Ποια είναι η μέση τιμή των τετραγώνων των διαφορών των παρατηρήσεων από τη μέση τιμή τους, δηλαδή η διακύμανση των βαρών κάθε ομάδας;

4) Εκφράζονται η διακύμανση και οι παρατηρήσεις με την ίδια μονάδα μέτρησης;
5) Ποια είναι η τετραγωνική ρίζα της διακύμανσης, δηλαδή η τυπική απόκλιση των βαρών κάθε ομάδας;
6) Εκφράζονται η τυπική απόκλιση και οι παρατηρήσεις με την ίδια μονάδα μέτρησης;
7) Να εκφράσετε το πηλίκο $\frac{\text{τυπική απόκλιση}}{\text{μέση τιμή}}$ ως ποσοστό για κάθε ομάδα βαρών. Ποια ομάδα έχει μικρότερο πηλίκο οπότε είναι περισσότερο ομοιογενής;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

*Ορισμοί για τα μέτρα θέσης ποσοτικών δεδομένων*

- Η **μέση τιμή $\bar{x}$** ενός συνόλου ν παρατηρήσεων ορίζεται, ως το άθροισμα των παρατηρήσεων διά του πλήθους τους.

$$\boxed{\bar{x} = \frac{t_1 + t_2 + t_3 + \ldots + t_\nu}{\nu}}$$

Αν για παράδειγμα οι βαθμοί ενός μαθητή σε δέκα μαθήματα είναι: 18, 17, 19, 17, 18, 16, 18, 17, 16, 18, τότε:

$$\bar{x} = \frac{18+17+19+17+18+16+18+17+16+18}{10} = \frac{174}{10} = 17{,}4$$

Αν οι τιμές των παρατηρήσεων είναι $x_i$ με αντίστοιχες συχνότητες $\nu_i$, τότε:

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

| $x_i$ | $\nu_i$ | $x_i \cdot \nu_i$ |
|---|---|---|
| 16 | 2 | 32 |
| 17 | 3 | 51 |
| 18 | 4 | 72 |
| 19 | 1 | 19 |
| Σύνολο | 10 | 174 |

Για το παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα έχουμε:

$$\bar{x} = \frac{16 \cdot 2 + 17 \cdot 3 + 18 \cdot 4 + 19 \cdot 1}{10} = \frac{174}{10} = 17{,}4$$

- Προκειμένου να υπολογίσουμε τη διάμεσο ενός δείγματος ν παρατηρήσεων, αρχικά διατάσσουμε τις παρατηρήσεις σε αύξουσα σειρά. Η **διάμεσος δ** ορίζεται ως:
  - η μεσαία παρατήρηση του διατεταγμένου δείγματος, όταν το ν είναι περιττός αριθμός,
  - το ημιάθροισμα των δύο μεσαίων παρατηρήσεων του διατεταγμένου δείγματος, όταν το ν είναι άρτιος αριθμός.

Για το παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα, αν τους διατάξουμε σε αύξουσα σειρά, έχουμε: 16, 16, 17, 17, 17, 18, 18, 18, 18, 19. Επομένως:

$$\delta = \frac{17+18}{2} = 17{,}5$$

Αν έχουμε τις παρατηρήσεις 4, 4, 5, 5, 6, 7, 39 σε αύξουσα σειρά, τότε η διάμεσος είναι δ = 5 ενώ η μέση τιμή είναι:

$$\bar{x} = \frac{4+4+5+5+6+7+39}{7} = \frac{70}{7} = 10$$

Παρατηρούμε ότι, στην παραπάνω περίπτωση, η διάμεσος επηρεάζεται λιγότερο από την ακραία παρατήρηση «39».

Είναι φανερό ότι το πολύ 50% των παρατηρήσεων είναι μικρότερες και το πολύ 50% είναι μεγαλύτερες από τη διάμεσο δ.

- Τα **τεταρτημόρια** συμβολίζονται με $Q_1$, $Q_2$ και $Q_3$.
  - Για το $Q_1$ έχουμε αριστερά το πολύ 25% και δεξιά το πολύ 75% των παρατηρήσεων.
  - Για το $Q_2$ έχουμε $Q_2$ = δ, δηλαδή συμπίπτει με τη διάμεσο.
  - Για το $Q_3$ έχουμε αριστερά το πολύ 75% και δεξιά το πολύ 25% των παρατηρήσεων.

Για τον ευκολότερο υπολογισμό των $Q_1$ και $Q_3$ έχουμε:

- Όταν το πλήθος των διατεταγμένων παρατηρήσεων είναι άρτιος αριθμός, διακρίνουμε το πρώτο μισό, του οποίου η διάμεσος είναι το $Q_1$ και το δεύτερο μισό, του οποίου η διάμεσος είναι το $Q_3$.
- Όταν το πλήθος των διατεταγμένων παρατηρήσεων είναι περιττός αριθμός, αφαιρούμε από το δείγμα τη διάμεσο και διακρίνουμε το πρώτο μισό, του οποίου η διάμεσος είναι το $Q_1$ και το δεύτερο μισό, του οποίου η διάμεσος είναι το $Q_3$.

Για το παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα είναι:

$$Q_1 = 17,\quad Q_2 = δ = 17{,}5,\quad Q_3 = 18$$

- **Επικρατούσα τιμή $M_0$** είναι η παρατήρηση με τη μεγαλύτερη συχνότητα. Στο παράδειγμα με τους βαθμούς του μαθητή στα δέκα μαθήματα, επικρατούσα τιμή είναι $M_0$ = 18. Η επικρατούσα τιμή μπορεί να μην είναι μοναδική. Όταν όλες οι παρατηρήσεις είναι διαφορετικές, τότε λέμε ότι δεν υπάρχει επικρατούσα τιμή.

***Ορισμοί για τα μέτρα διασποράς ποσοτικών δεδομένων***

- Το **εύρος R** είναι το απλούστερο από τα μέτρα διασποράς και ορίζεται ως η διαφορά της ελάχιστης $x_{min}$ παρατήρησης από τη μέγιστη $x_{max}$ παρατήρηση. Δηλαδή:

$$\boxed{R = \text{μέγιστη παρατήρηση} - \text{ελάχιστη παρατήρηση} = x_{max} - x_{min}}$$

- Το **ενδοτεταρτημοριακό εύρος Q** είναι η διαφορά του πρώτου τεταρτημορίου $Q_1$ από το τρίτο τεταρτημόριο $Q_3$. Δηλαδή:

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

Όσες παρατηρήσεις βρίσκονται έξω από το διάστημα $[Q_1 - 1{,}5 \cdot Q\ ,\ Q_3 + 1{,}5 \cdot Q]$ ονομάζονται **ακραίες τιμές**. Η διάμεσος είναι ένα μέτρο θέσης το οποίο δεν επηρεάζεται από το μέγεθος των ακραίων τιμών.

Το εύρος και το ενδοτεταρτημοριακό εύρος είναι πολύ απλά μέτρα διασποράς και δεν εκφράζουν την απόκλιση που έχει κάθε παρατήρηση από το «κέντρο» των παρατηρήσεων. Για τον λόγο αυτό χρησιμοποιούμε και τα επόμενα μέτρα διασποράς.

- Η **διακύμανση $s^2$** ορίζεται ως η μέση τιμή των τετραγώνων των διαφορών της μέσης τιμής των παρατηρήσεων από τις παρατηρήσεις. Δηλαδή:

$$\boxed{s^2 = \frac{(t_1 - \bar{x})^2 + (t_2 - \bar{x})^2 + (t_3 - \bar{x})^2 + \ldots + (t_\nu - \bar{x})^2}{\nu}}$$

Για παράδειγμα, αν οι βαθμοί έξι μαθητών στα Μαθηματικά είναι: 18, 15, 18, 18, 15, 18, τότε:

$$\bar{x} = \frac{18+15+18+18+15+18}{6} = \frac{102}{6} = 17 \text{ και}$$

$$s^2 = \frac{(18-17)^2 + (15-17)^2 + (18-17)^2 + (18-17)^2 + (15-17)^2 + (18-17)^2}{6} = \frac{12}{6} = 2$$

Αν οι τιμές των παρατηρήσεων είναι $x_i$ με αντίστοιχες συχνότητες $\nu_i$, τότε:

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

Για το παράδειγμα των βαθμών των έξι μαθητών στα Μαθηματικά, σχετικός είναι ο επόμενος πίνακας:

| $x_i$ | $\nu_i$ | $x_i \cdot \nu_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^2 \cdot \nu_i$ |
|---|---|---|---|---|---|
| 15 | 2 | 30 | -2 | 4 | 8 |
| 18 | 4 | 72 | 1 | 1 | 4 |
| Σύνολο | 6 | 102 | | | 12 |

$$\bar{x} = \frac{102}{6} = 17 \quad \text{και} \quad s^2 = \frac{12}{6} = 2$$

Η διακύμανση δεν εκφράζεται με τις μονάδες που εκφράζονται οι παρατηρήσεις.

- Η **τυπική απόκλιση s** είναι: $\boxed{s = \sqrt{s^2}}$ Στο παραπάνω παράδειγμα είναι: $s = \sqrt{2} \approx 1{,}41$. Η τυπική απόκλιση έχει το πλεονέκτημα ότι εκφράζεται με τις μονάδες, που εκφράζονται οι παρατηρήσεις.

***Θηκόγραμμα***

- Έστω ότι έχουμε τον αριθμό των εργαζομένων σε 20 βιοτεχνίες.

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 14 | 25 | 7 | 31 | 8 | 12 | 19 | 10 | 24 | 9 | 13 | 5 | 28 | 24 | 19 | 26 | 51 | 68 | 14 |

Κατατάσσουμε τις παρατηρήσεις σε αύξουσα σειρά:

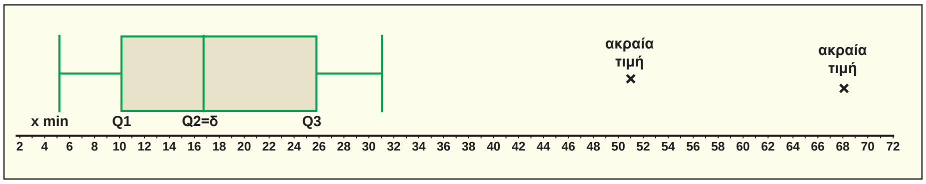
5, 7, 8, 9, 10, 10, 12, 13, 14, 14, 19, 19, 24, 24, 25, 26, 28, 31, 51, 68

Είναι: $Q_1 = \dfrac{10+10}{2} = 10,\ Q_2 = \delta = \dfrac{14+19}{2} = 16{,}5,\ \ Q_3 = \dfrac{25+26}{2} = 25{,}5$

Το ενδοτεταρτημοριακό εύρος είναι: Q = 25,5 - 10 = 15,5 και 1,5 · Q = 23,25

Επομένως $\left[Q_1 - 1{,}5 \cdot Q\,,\, Q_3 + 1{,}5 \cdot Q\right] = \left[-13{,}25\,,\, 48{,}75\right]$. Οπότε ακραίες παρατηρήσεις είναι το 51 και το 68.

Τα μέτρα αυτά απεικονίζονται στο επόμενο διάγραμμα που λέγεται **θηκόγραμμα**.




Από τα μέσα των πλευρών, που παριστάνουν το πρώτο και το τρίτο τεταρτημόριο (δηλαδή τα $Q_1$ και $Q_3$ αντιστοίχως) φέρουμε ευθύγραμμα τμήματα με μήκος που προσδιορίζεται ως εξής:

Η μικρότερη παρατήρηση που είναι μεγαλύτερη του $Q_1$-1,5·Q είναι το 5, επομένως το αριστερό άκρο του τμήματος που θα φέρουμε από το $Q_1$ είναι το 5 και δεν υπάρχουν παρατηρήσεις μικρότερες του $Q_1$-1,5·Q για να σημειωθούν χωριστά.

Η μεγαλύτερη παρατήρηση που είναι μικρότερη του $Q_3$+1,5·Q είναι η 31, επομένως το δεξί άκρο του τμήματος που θα φέρουμε από το $Q_3$ είναι το 31 και θα σημειώσουμε χωριστά τις ακραίες τιμές 51 και 68 οι οποίες είναι μεγαλύτερες του $Q_3$+1,5·Q.

***Συντελεστής μεταβλητότητας***

- Δυο όμιλοι επιχειρήσεων, που ο καθένας αποτελείται από 5 επιχειρήσεις, είχαν ετήσιες δαπάνες για το οικονομικό έτος 2017 τα ποσά, σε χιλιάδες ευρώ, που αναγράφονται στον παρακάτω πίνακα.

| | | | | | |
|---|---|---|---|---|---|
| **Όμιλος Α** | 200 | 250 | 300 | 300 | 350 |
| **Όμιλος Β** | 5.000 | 5.050 | 5.100 | 5.100 | 5.150 |

Οι μέσες τιμές των δαπανών για κάθε όμιλο είναι:

$$\bar{x}_A = \frac{200+250+300+300+350}{5} = 280 \text{ χιλιάδες ευρώ.}$$

$$\bar{x}_B = \frac{5.000+5.050+5.100+5.100+5.100+5.150}{5} = 5.080 \text{ χιλιάδες ευρώ.}$$

Οι διακυμάνσεις των δαπανών για κάθε όμιλο είναι:

$$s_A^2 = \frac{(200-280)^2+(250-280)^2+(300-280)^2+(300-280)^2+(350-280)^2}{5}$$

$$s_A^2 = 2.600$$

$$s_B^2 = \frac{(5.000-5.080)^2+(5.050-5.080)^2+(5.100-5.080)^2+(5.100-5.080)^2+(5.150-5.080)^2}{5}$$

$$s_B^2 = 2.600$$

Οι τυπικές αποκλίσεις των δαπανών για κάθε όμιλο είναι ίσες με:

$$s_A = s_B = \sqrt{2.600} \approx 51 \text{ χιλιάδες ευρώ.}$$

Μπορούμε άραγε να ισχυριστούμε ότι οι 51 χιλιάδες ευρώ των ίσων τυπικών αποκλίσεων έχουν την ίδια βαρύτητα για τους ομίλους A και B, που έχουν διαφορετικές μέσες τιμές $\bar{x}_A = 280$ χιλιάδες ευρώ και $\bar{x}_B = 5.080$ χιλιάδες ευρώ; Ένας τέτοιος ισχυρισμός θα ήταν λανθασμένος, γιατί η σχέση τιμών της μεταβλητής «ετήσιες δαπάνες» με την τιμή της τυπικής απόκλισης, είναι διαφορετικές για τις επιχειρήσεις των δυο ομίλων.

Ένα μέτρο, που μας βοηθάει να αντιμετωπίσουμε τέτοια προβλήματα και μας δίνει τη δυνατότητα συγκρίσεων, είναι ο **συντελεστής μεταβλητότητας CV**, ο οποίος είναι καθαρός αριθμός και εκφράζεται συνήθως ως ποσοστό. Ορίζεται ως ακολούθως:

$$\boxed{CV = \frac{s}{|\bar{x}|}}$$

Στο παράδειγμά μας ο συντελεστής μεταβλητότητας για τις δυο επιχειρήσεις είναι:

$$CV_A = \frac{51}{280} \approx 0,1821 = 18,21\% \text{ και } CV_B = \frac{51}{5.080} \approx 0,01004 \approx 1,00\%$$

Όσο μικρότερο είναι το ποσοστό αυτό, τόσο περισσότερη ομοιογένεια υπάρχει στις τιμές της μεταβλητής που μας ενδιαφέρει. Στο παράδειγμά μας είναι $CV_A > CV_B$, επομένως η διασπορά των τιμών στον όμιλο A σε σχέση με τη μέση τιμή είναι μεγαλύτερη από την αντίστοιχη διασπορά στον όμιλο B.

## Εφαρμογές

### Εφαρμογή 1

Μία ομάδα δέκα μαθητών μέτρησε το μήκος ενός θρανίου με ακρίβεια δέκατου του εκατοστού του μέτρου (χιλιοστού του μέτρου). Οι μαθητές χρησιμοποίησαν την ίδια μετροταινία και κατέγραψαν τις τιμές των δέκα μετρήσεων, όπως φαίνεται στον παρακάτω πίνακα.

| Μήκος σε cm | 120,2 | 120,1 | 120,1 | 119,8 | 119,7 | 120,3 | 120,2 | 119,9 | 120,1 | 119,8 |
|---|---|---|---|---|---|---|---|---|---|---|

1) Να αναφέρετε δύο λόγους που, κατά τη γνώμη σας, δικαιολογούν τις διαφορές στις μετρήσεις του θρανίου.
2) Πώς υπολογίζουμε τη μέση τιμή των μετρήσεων;
3) Ποια είναι η τιμή του μήκους του θρανίου που θα χρησιμοποιήσουμε;
4) Για ποιον λόγο είναι χρήσιμος ο υπολογισμός της μέσης τιμής πολλών μετρήσεων;
5) Δικαιολογήστε με παράδειγμα -που θα δημιουργήσετε από τις τιμές του πίνακα- γιατί ζητήθηκε από τους μαθητές να κάνουν δέκα μετρήσεις και όχι λιγότερες (π.χ. τρεις μετρήσεις).

**<u>Λύση</u>**

1) Δύο λόγοι που δικαιολογούν τις διαφορές μπορεί να είναι:
   1) Η αρχή της μετροταινίας δε συμπίπτει με την αρχή της μετρούμενης απόστασης
   2) Η μετροταινία δεν ακολουθεί ευθεία και παράλληλη γραμμή προς τη μετρούμενη απόσταση.
2) Για τη μέση τιμή των μετρήσεων έχουμε:

$$\bar{x} = \frac{120{,}2+120{,}1+120{,}1+119{,}8+119{,}7+120{,}3+120{,}2+119{,}9+120{,}1+119{,}8}{10}$$

$$= \frac{1200{,}2}{10} = 120{,}02$$

3) Για την άγνωστη τιμή του μήκους του θρανίου, μπορούμε να χρησιμοποιήσουμε τη μέση τιμή $\bar{x} = 120{,}02$ cm του δείγματος.
4) Θέλουμε τη μέση τιμή πολλών μετρήσεων, δηλαδή θέλουμε μεγάλο μέγεθος δείγματος, ώστε το αποτέλεσμα να έχει μεγαλύτερη πιθανότητα να προσεγγίσει ικανοποιητικά το πραγματικό μήκος του θρανίου.
5) Αν για παράδειγμα πάρουμε τη μέση τιμή των τριών πρώτων μετρήσεων έχουμε:

$$\bar{x}' = \frac{120{,}2+120{,}1+120{,}1}{3} = \frac{360{,}4}{3} = 120{,}133\ldots \text{ cm}$$

Αυτή είναι η μέση τιμή ενός άλλου δείγματος. Το μήκος όμως του θρανίου είναι μοναδικό. Το μεγαλύτερο μέγεθος του δείγματος μας δίνει τη δυνατότητα ακριβέστερης εκτίμησης του πραγματικού μήκους.

## Εφαρμογή 2

Τα αποτελέσματα του SAT τεστ στις ΗΠΑ, για μαθητές και μαθήτριες που τελειώνοντας το σχολείο προτίθενται να συνεχίσουν τις σπουδές τους, έδειξαν το 2019 τα εξής :

Πήραν μέρος 2.220.087 μαθητές και μαθήτριες και η μεγαλύτερη δυνατή επίδοση (χωρίς κανένα λάθος/καμία απώλεια βαθμών) ήταν 1.600. Η μέση επίδοση του παραπάνω πληθυσμού στο τεστ ήταν 1.059.

Στο πλαίσιο μίας διερευνητικής εργασίας που έγινε σε εννέα διαφορετικές τάξεις, δέκα ομάδες μαθητών/τριών από κάθε τάξη συγκέντρωσαν η καθεμία τις επιδόσεις ενός διαφορετικού δείγματος μαθητών και μαθητριών πλήθους 100 (προσπαθώντας κάθε δείγμα να είναι, κατά το δυνατόν, αντιπροσωπευτικό του πληθυσμού) και υπολόγισαν τη μέση επίδοση του δείγματος.

Τα αποτελέσματα της δουλειάς των ομάδων φαίνονται στον παρακάτω πίνακα:

| | Ομάδες | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Τάξεις** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **1** | 1063 | 1059 | 1047 | 1056 | 1057 | 1040 | 1038 | 1076 | 1053 | 1038 |
| **2** | 1071 | 1039 | 1062 | 1056 | 1071 | 1054 | 1043 | 1067 | 1078 | 1048 |
| **3** | 1061 | 1074 | 1068 | 1060 | 1076 | 1047 | 1056 | 1061 | 1061 | 1057 |
| **4** | 1040 | 1072 | 1039 | 1076 | 1043 | 1063 | 1072 | 1067 | 1061 | 1056 |
| **5** | 1051 | 1071 | 1051 | 1072 | 1049 | 1049 | 1080 | 1055 | 1049 | 1044 |
| **6** | 1072 | 1056 | 1070 | 1049 | 1045 | 1051 | 1061 | 1049 | 1076 | 1071 |
| **7** | 1065 | 1058 | 1080 | 1056 | 1079 | 1040 | 1068 | 1075 | 1053 | 1072 |
| **8** | 1077 | 1038 | 1069 | 1065 | 1054 | 1055 | 1054 | 1065 | 1060 | 1072 |
| **9** | 1069 | 1076 | 1069 | 1079 | 1041 | 1051 | 1048 | 1064 | 1048 | 1063 |

Ο καθηγητής της τάξης 1 αποφάσισε να υπολογίσει μία νέα μέση επίδοση συνολικά, με χρήση και των δέκα που υπολόγισαν οι ομάδες των μαθητών/τριών του.

Πώς σχολιάζετε την ενέργεια αυτή;

Τι παρατηρείτε σχετικά με το αποτέλεσμα που προκύπτει;

(Πηγή: SAT Suite of Assessments Annual Report 2019, ανακτήθηκε από https://research.collegeboard.org/programs/sat/data/2019-sat-suite-annual-report, τελευταία προσπέλαση, 28/11/2019)

**Λύση**

Η ενέργεια του καθηγητή έχει νόημα καθώς το δείγμα συνεχίζει να είναι αντιπροσωπευτικό, με τη διαφορά ότι έχει μέγεθος 1000 (δεκαπλάσιο).
Το αποτέλεσμα προκύπτει ως εξής:

$$\bar{x}_1 = \frac{1063+1059+1047+1056+1057+1040+1038+1076+1053+1038}{10} = 1052,7$$

Με όμοιο τρόπο έχουμε (για τις υπόλοιπες τάξεις):

| $\bar{x}_2$ | $\bar{x}_3$ | $\bar{x}_4$ | $\bar{x}_5$ | $\bar{x}_6$ | $\bar{x}_7$ | $\bar{x}_8$ | $\bar{x}_9$ |
|---|---|---|---|---|---|---|---|
| 1058,9 | 1062,1 | 1058,9 | 1057,1 | 1060 | 1064,6 | 1060,9 | 1060,8 |

Όπως παρατηρούμε, οι νέες μέσες επιδόσεις των νέων δειγμάτων με πλήθος 1000 αντί για 100 είναι «συνολικά πιο κοντά» στη μέση επίδοση του πληθυσμού (που είναι 1059), από τις μέσες επιδόσεις των δειγμάτων με πλήθος 100.
Επίσης, αν θεωρήσουμε όλα τα δείγματα σαν ένα δείγμα πλήθους 9.000, η μέση επίδοση αυτού του ακόμα μεγαλύτερου δείγματος θα είναι 1059,556.

***Στατιστικά μέτρα δείγματος και πληθυσμού***

**Γενικότερα**

Σε ένα δείγμα παρατηρήσεων το μέσο $\bar{x}$ τον λέμε δειγματικό μέσο και την τυπική απόκλιση s τη λέμε δειγματική τυπική απόκλιση και μπορούμε να τα υπολογίσουμε. Εξαρτώνται βέβαια κάθε φορά από την επιλογή του δείγματος. Όμως, μέσω του δείγματος, τον μέσο μ του πληθυσμού και την τυπική του απόκλιση σ μπορούμε μόνο να τα εκτιμήσουμε και όχι να τα υπολογίσουμε.

## Εφαρμογή 3

Η μέση τιμή των μηνιαίων μισθών των υπαλλήλων μιας εταιρείας είναι 1.200 ευρώ και η τυπική απόκλιση 100 ευρώ.

1) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 50 ευρώ, τότε ποια μεταβολή επέρχεται στη μέση τιμή, στην τυπική απόκλιση και στον συντελεστή μεταβλητότητας των μισθών;

2) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 5%, τότε ποια μεταβολή επέρχεται στη μέση τιμή, στην τυπική απόκλιση και στον συντελεστή μεταβλητότητας των μισθών;

**Λύση**

Έστω ότι $t_1, t_2, t_3, \ldots, t_\nu$ είναι οι μηνιαίοι μισθοί των υπαλλήλων της εταιρείας. Τότε:

✓ $$\bar{x} = \frac{t_1+t_2+t_3+\ldots+t_\nu}{\nu} = 1.200$$

✓ $$s = \sqrt{\frac{(t_1-1.200)^2+(t_2-1.200)^2+\ldots+(t_\nu-1.200)^2}{\nu}} = 100$$

✓ $CV=\dfrac{s}{|\bar{x}|}=\dfrac{100}{1.200}\approx 8{,}33\%$

1) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 50 ευρώ, τότε:

$$\bar{y}=\frac{(t_1-50)+(t_2-50)+...+(t_\nu-50)}{\nu}=\frac{t_1+t_2+...+t_\nu}{\nu}-\frac{50\cdot\nu}{\nu}=1.200-50=1.150$$

$$s_y=\sqrt{\frac{(t_1-50-1.150)^2+...+(t_\nu-50-1.150)^2}{\nu}}=\sqrt{\frac{(t_1-1.200)^2+...+(t_\nu-1.200)^2}{\nu}}=s=100$$

$$CV_y=\frac{s_y}{|\bar{y}|}=\frac{100}{1.150}\approx 8{,}7\%$$

Επομένως:
✓ Η μέση τιμή $\bar{x}=1.200$€ μειώνεται κατά 50€ και γίνεται $\bar{y}=1.150$ €
✓ Η τυπική απόκλιση δε μεταβάλλεται.
✓ Ο συντελεστής μεταβλητότητας CV = 8,33% αυξάνεται και γίνεται $CV_y$ = 8,7%

2) Αν γίνει μια κράτηση σε κάθε μηνιαίο μισθό κατά 5%, τότε όλοι οι μηνιαίοι μισθοί πολλαπλασιάζονται με 0,95. Έτσι έχουμε:

$$\bar{z}=\frac{0{,}95\cdot t_1+0{,}95\cdot t_2+...+0{,}95\cdot t_\nu}{\nu}=0{,}95\cdot\frac{t_1+t_2+...+t_\nu}{\nu}=0{,}95\cdot 1200$$

$$\bar{z}=1.140$$

$$s_z=\sqrt{\frac{(0{,}95\cdot t_1-0{,}95\cdot\bar{x})^2+(0{,}95\cdot t_2-0{,}95\cdot\bar{x})^2+...+(0{,}95\cdot t_\nu-0{,}95\cdot\bar{x})^2}{\nu}}$$

$$=\sqrt{0{,}95^2\cdot s^2}=|0{,}95|\cdot s=0{,}95\cdot 100=95$$

$$CV_z=\frac{s_z}{|\bar{z}|}=\frac{0{,}95\cdot s}{0{,}95\cdot\bar{x}}=CV=8{,}33\%$$

Επομένως:
✓ Η μέση τιμή $\bar{x}=1.200$€ μειώνεται κατά 5% και γίνεται $\bar{z}=1.140$€
✓ Η τυπική απόκλιση s = 100 € μειώνεται κατά 5% και γίνεται $s_z$ = 95 €
✓ Ο συντελεστής μεταβλητότητας δε μεταβάλλεται.

*Έστω ότι έχουμε τις τιμές $x_i$, i = 1,2,...,ν, μιας ποσοτικής μεταβλητής X με μέση τιμή $\bar{x}$ και τυπική απόκλιση $s_x$. Τότε για τη μέση τιμή $\bar{y}$ και την τυπική απόκλιση $s_y$ των τιμών της μεταβλητής Y για τις οποίες ισχύει $y_i=\alpha x_i+\beta$, i=1,2,...,ν και $\alpha,\beta\in R$ είναι: $\bar{y}=\alpha\bar{x}+\beta$ και $s_y=|\alpha|\cdot s_x$*

## Εφαρμογή 4

Στην εθνική οδό Αθηνών - Λαμίας τα αυτοκίνητα τρέχουν με σταθερή ταχύτητα. Εμείς, με το δικό μας αυτοκίνητο, τρέχουμε με 100 Km/h και προσπερνάμε σε πλήθος όσα ακριβώς μας προσπερνούν. Να βρείτε τη διάμεσο των ταχυτήτων του δικού μας αυτοκινήτου και όλων όσων μας προσπέρασαν ή προσπεράσαμε στο ταξίδι μας αυτό.

**Λύση**

Η διάμεσος είναι δ = 100 Km/h διότι το πολύ 50% των ταχυτήτων είναι μικρότερες και το πολύ 50% είναι μεγαλύτερες από 100 Km/h.

## Εφαρμογή 5

Σε μια κάλπη υπάρχουν άσπρες, μαύρες, κόκκινες και πράσινες μπάλες σε αναλογία 10%, 20%, 30% και 40% αντίστοιχα. Μια άσπρη μπάλα έχει βάρος 10 gr, μια μαύρη 11 gr, μια κόκκινη 12 gr και μια πράσινη 13 gr. Να βρείτε τη μέση τιμή του βάρους τους, αν γνωρίζουμε ότι στην κάλπη υπάρχουν:

1) 10 μπάλες.
2) 20 μπάλες.
3) Δε γνωρίζουμε πόσες μπάλες υπάρχουν στην κάλπη.

**Λύση**

| $x_i$ | $v_i$ | $f_i$ |
|---|---|---|
| 10 | 1 | 0,10 |
| 11 | 2 | 0,20 |
| 12 | 3 | 0,30 |
| 13 | 4 | 0,40 |
| Σύνολο | 10 | 1,00 |

1) Με βάση την αναλογία της εκφώνησης, στις 10 μπάλες θα έχουμε 1 άσπρη, 2 μαύρες, 3 κόκκινες και 4 πράσινες. Αν οι τιμές των βαρών είναι οι παρατηρήσεις $x_i$ με αντίστοιχες συχνότητες $v_i$ και σχετικές συχνότητες $f_i$, τότε:

$$\bar{x} = \frac{1\cdot 10 + 2\cdot 11 + 3\cdot 12 + 4\cdot 13}{10}$$
$$= 0{,}10\cdot 10 + 0{,}20\cdot 11 + 0{,}30\cdot 12 + 0{,}40\cdot 13 = 12\text{ gr}$$

| $x_i$ | $v_i$ | $f_i$ |
|---|---|---|
| 10 | 2 | 0,10 |
| 11 | 4 | 0,20 |
| 12 | 6 | 0,30 |
| 13 | 8 | 0,40 |
| Σύνολο | 20 | 1,00 |

2) Με βάση την αναλογία της εκφώνησης, στις 20 μπάλες θα έχουμε 2 άσπρες, 4 μαύρες, 6 κόκκινες και 8 πράσινες. Αν οι τιμές των βαρών είναι οι παρατηρήσεις $x_i$ με αντίστοιχες συχνότητες $v_i$ και σχετικές συχνότητες $f_i$, τότε:

$$\bar{x} = \frac{2\cdot 10 + 4\cdot 11 + 6\cdot 12 + 8\cdot 13}{20}$$
$$= 0{,}10\cdot 10 + 0{,}20\cdot 11 + 0{,}30\cdot 12 + 0{,}40\cdot 13 = 12\text{ gr}$$

3) Παρατηρώντας προσεκτικά την παραπάνω διαδικασία καταλήγουμε ότι ανεξαρτήτως του πλήθους από τις μπάλες έχουμε:

$$\bar{x} = f_1\cdot x_1 + f_2\cdot x_2 + f_3\cdot x_3 + f_4\cdot x_4 = 0{,}10\cdot 10 + 0{,}20\cdot 11 + 0{,}30\cdot 12 + 0{,}40\cdot 13 =$$
$$= 12\text{ gr}$$

*Έστω ότι έχουμε τις τιμές $x_i$, i = 1, 2,…,ν με σχετικές συχνότητες $f_i$, i = 1, 2,…,ν. Τότε για τη μέση τιμή $\bar{x}$ είναι:*

$$\bar{x} = f_1 \cdot x_1 + f_2 \cdot x_2 + \ldots + f_\nu \cdot x_\nu$$

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Οι βαθμοί του Αντρέα σε 4 διαγωνίσματα στα Μαθηματικά ήταν 15, 18, 18, 17. Για τα ίδια διαγωνίσματα, ο Βασίλης είχε πάρει 2 μονάδες περισσότερες σε κάθε διαγώνισμα από τον Αντρέα, ενώ ο Γιάννης είχε πάρει 4 μονάδες λιγότερες από τον Αντρέα σε κάθε διαγώνισμα. Να βρείτε τη μέση τιμή των βαθμών του κάθε παιδιού.

**2)** Αν ο μέσος όρος του μηνιαίου μισθού των υπαλλήλων ενός εργοστασίου πέρυσι ήταν 850€ και φέτος σε κάθε υπάλληλο δοθεί αύξηση 50€, να βρείτε τον νέο μέσο όρο των μισθών.

**3)** Καθεμία από τις παρακάτω λίστες δεδομένων έχουν μέση τιμή 50.

(I) 0, 20, 40, 50, 60, 80, 100
(II) 0, 48, 49, 50, 51, 52, 100
(III) 0, 1, 2, 50, 98, 99, 100

1) Μπορεί να χρησιμοποιηθεί το εύρος για σύγκριση της μεταβλητότητας των δεδομένων αυτών;
2) Χωρίς να γίνουν οι πράξεις, να βρείτε σε ποια λίστα υπάρχει μεγαλύτερη και σε ποια μικρότερη διασπορά παρατηρήσεων.

**4)** Να υπολογίσετε τη μέση τιμή και τη διάμεσο για τα παρακάτω δείγματα δεδομένων και να σχολιάσετε τα αποτελέσματα:

α) 1, 2, 6   β) 2, 4, 12   γ) 11, 12, 16   δ) 12, 14, 22

**5)** Να υπολογίσετε τη διακύμανση και την τυπική απόκλιση για καθεμιά από τις παρακάτω λίστες δεδομένων. Συγκρίνοντας τα δεδομένα και τα αποτελέσματα, τι συμπέρασμα βγάζετε;

α) 1, 3, 4, 5, 7   β) 3, 9, 12, 15, 21   γ) 6, 8, 9, 10, 12   δ) -1, -3, -4, -5, -7

**6)** Η βαθμολογία 16 μαθητών σε ένα διαγώνισμα ήταν: 8, 15, 13, 20, 9, 13, 17, 19, 20, 9, 10, 10, 15, 13, 14, 17. Να υπολογίσετε:

1) Τα τρία μέτρα θέσης, μέση τιμή, διάμεσο και επικρατούσα τιμή.
2) Το εύρος, την τυπική απόκλιση και τον συντελεστή μεταβλητότητας.

**7)** Η μέση τιμή ηλικίας των υπαλλήλων μιας εταιρείας είναι 32 χρόνια. Ποια θα είναι η μέση τιμή ηλικίας των ίδιων υπαλλήλων ύστερα από τρία χρόνια;

**8)** Οι βαθμοί στα Μαθηματικά 20 μαθητών της Β΄ τάξης ενός Λυκείου είναι:

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | 14 | 15 | 13 | 17 | 15 | 16 | 14 | 18 | 15 | 17 | 13 | 19 | 15 | 16 | 12 | 16 | 18 | 13 | 14 |

1) Να βρείτε τη μέση τιμή και την επικρατούσα τιμή.
2) Να βρείτε τη διάμεσο.
3) Να βρείτε το πρώτο και το τρίτο τεταρτημόριο.
4) Να σχεδιάσετε το θηκόγραμμα.

# Πρόσθετο Υλικό

## Σύνθετες ασκήσεις

**1)** Η μέση επίδοση 17 αγοριών και 13 κοριτσιών στο μάθημα των Μαθηματικών μιας τάξης είναι 16,8. Η μέση επίδοση των κοριτσιών είναι 15,6. Να βρείτε τη μέση επίδοση των αγοριών.

**2)** Σε ένα Λύκειο υπάρχουν 500 μαθητές. Η Α΄ τάξη έχει 200 μαθητές με μέσο όρο ηλικίας 15,7 χρόνια, ενώ η Β΄ τάξη έχει 180 μαθητές με μέσο όρο ηλικίας 16,9 χρόνια. Οι υπόλοιποι μαθητές της Γ΄ τάξης έχουν μέσο όρο ηλικίας 17,7 χρόνια. Να υπολογίσετε τον μέσο όρο ηλικίας όλων των μαθητών του σχολείου.

**3)** Η μέση τιμή 40 παρατηρήσεων είναι 20. Αν από αυτές οι 7 μειώνονται κατά 2 και οι 9 αυξάνονται κατά 6, να βρεθεί η νέα μέση τιμή.

**4)** Η τυπική απόκλιση ενός δείγματος με ν παρατηρήσεις είναι ίση με μηδέν. Ποιο συμπέρασμα προκύπτει για τις παρατηρήσεις;

**5)** Η μέση τιμή 6 διαδοχικών ακέραιων αριθμών είναι 7,5. Να υπολογίσετε την τυπική τους απόκλιση.

**6)** Αν σε μία τάξη ο μέσος όρος της βαθμολογίας $\nu_1$ αγοριών είναι $\bar{x}$ και ο μέσος όρος της βαθμολογίας $\nu_2$ κοριτσιών είναι $\bar{y}$, να αποδείξετε ότι ο μέσος όρος της βαθμολογίας όλων των παιδιών της τάξης είναι $\bar{z} = \dfrac{\nu_1\bar{x} + \nu_2\bar{y}}{\nu_1 + \nu_2}$.

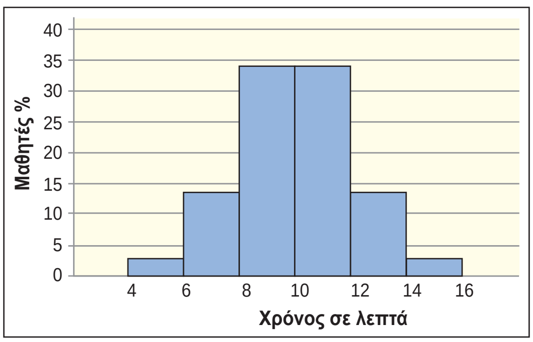
**7)** Να αποδείξετε ότι η μέση τιμή είναι το άθροισμα των γινομένων των τιμών της μεταβλητής επί τις αντίστοιχες σχετικές συχνότητες.

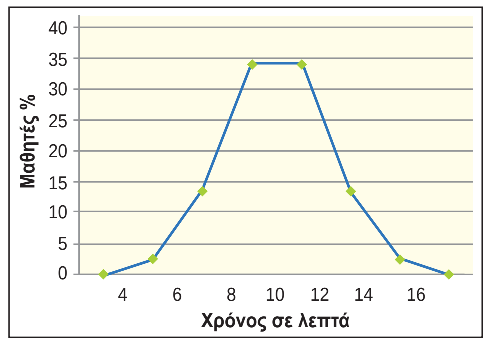
**8)** Σε ένα εργοστάσιο με 100 εργαζόμενους η μέση τιμή των αμοιβών τους είναι 900€. Οι 40 από αυτούς πληρώνονται με μισθό μικρότερο της μέσης τιμής και οι μισθοί τους έχουν μέση τιμή 800€. Αν οι αποδοχές των εργαζομένων με μισθό μικρότερο της μέσης τιμής αυξηθούν και γίνουν όσο η μέση τιμή, τότε ποια θα είναι η νέα μέση τιμή των αμοιβών των 100 εργαζομένων;

# ΕΝΟΤΗΤΑ 2.4 : ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

## Διερεύνηση

Το παρακάτω ιστόγραμμα και το αντίστοιχο πολύγωνο σχετικών συχνοτήτων παρουσιάζουν τον χρόνο, σε λεπτά, που χρειάζονται οι μαθητές για να πάνε από το σπίτι τους στο σχολείο. Η μέση τιμή είναι 10 λεπτά και η τυπική απόκλιση 2 λεπτά.



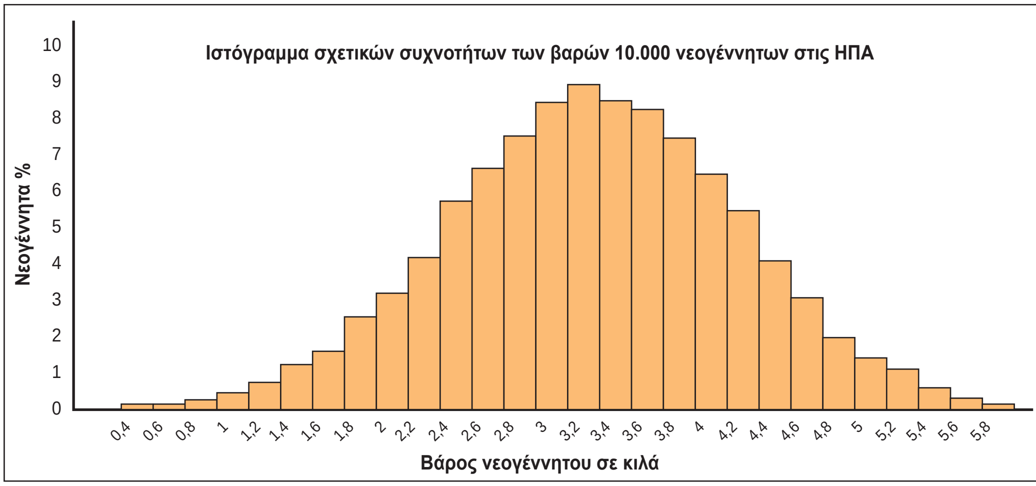




1) Τι έχετε να παρατηρήσετε για τη συμμετρία της κατανομής;
2) Τι ποσοστό μαθητών χρειάζονται το πολύ 10 λεπτά και πόσοι τουλάχιστον 10 λεπτά για να πάνε στο σχολείο;
3) Τι ποσοστό μαθητών χρειάζονται από 8 εως 12 λεπτά για να πάνε στο σχολείο;
4) Τι ποσοστό μαθητών χρειάζονται τουλάχιστον 8 λεπτά για να πάνε στο σχολείο;

## Βασικές μαθηματικές έννοιες - Ιδέες - Διεργασίες

***Κανονική κατανομή***

Στο παρακάτω ιστόγραμμα σχετικών συχνοτήτων επί τοις εκατό (%) παριστάνεται η κατανομή 10.000 νεογέννητων στις ΗΠΑ, ως προς το βάρος τους (σε κιλά). Τα νεογέννητα κατανέμονται σε κλάσεις ίσου πλάτους. Από αυτό το ιστόγραμμα αντλούμε αρκετές πληροφορίες, όπως π.χ. ότι η κλάση με τα περισσότερα νεογέννητα είναι από 3,2 έως 3,4 κιλά, η οποία περιέχει γύρω στο 9% του δείγματος, δηλαδή 900 νεογέννητα περίπου.




Ακολουθεί το αντίστοιχο πολύγωνο σχετικών συχνοτήτων (%) της μεταβλητής «βάρος νεογέννητου στις ΗΠΑ», για το ίδιο δείγμα των 10.000. Παρατηρούμε ότι η μορφή του είναι σχεδόν συμμετρική ως προς μία κατακόρυφη ευθεία, η οποία είναι κάθετη στον άξονα των βαρών, κοντά στα 3,2 κιλά. Από τη μορφή του πολυγώνου συχνοτήτων καταλαβαίνουμε ότι:

- σχεδόν το 50% των 10.000 νεογέννητων έχει βάρος μικρότερο των 3,2 κιλών και σχεδόν το 50% έχει βάρος μεγαλύτερο των 3,2 κιλών,
- η κατανομή των βαρών «γύρω από τα 3,2 κιλά» γίνεται με συμμετρικό τρόπο,
- καθώς κατευθυνόμαστε από τις ακραίες τιμές βάρους προς τα 3,2 κιλά (δηλαδή προς το «κέντρο» της κατανομής), τα νεογέννητα που κατανέμονται στις κλάσεις είναι περισσότερα.

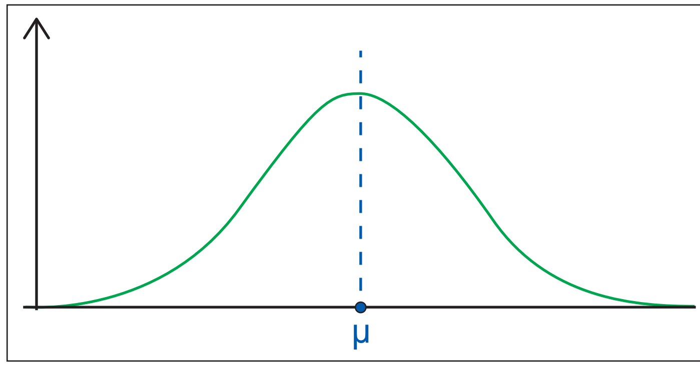

Στο ιστόγραμμα σχετικών συχνοτήτων επί τοις εκατό (%) που ακολουθεί, παριστάνεται η κατανομή 10.000 ενηλίκων, ως προς τη μέτρηση της διαστολικής αρτηριακής τους πίεσης σε mmHg (χιλιοστά στήλης υδραργύρου), σε κατάσταση ηρεμίας. Κι εδώ, η κατανομή γίνεται σε κλάσεις ίσου πλάτους 5 mmHg. Τα ιστογράμματα των δύο παραδειγμάτων (βάρους νεογέννητων και αρτηριακής πίεσης ενηλίκων) έχουν παρόμοια μορφή.



Η ομοιότητα των δυο προηγούμενων κατανομών φαίνεται, επίσης, και στα πολύγωνα σχετικών συχνοτήτων (%). Στην περίπτωση της μεταβλητής «διαστολική αρτηριακή πίεση ενηλίκων στην κεντρική Ευρώπη», στο δείγμα εμφανίζεται παρόμοια συμμετρία ως προς την ευθεία που είναι κάθετη στον άξονα των τιμών πίεσης, στα 85 mmHg περίπου. Και εδώ, μπορούμε να κάνουμε παρόμοιες παρατηρήσεις με εκείνες του παραδείγματος με τα βάρη νεογέννητων.

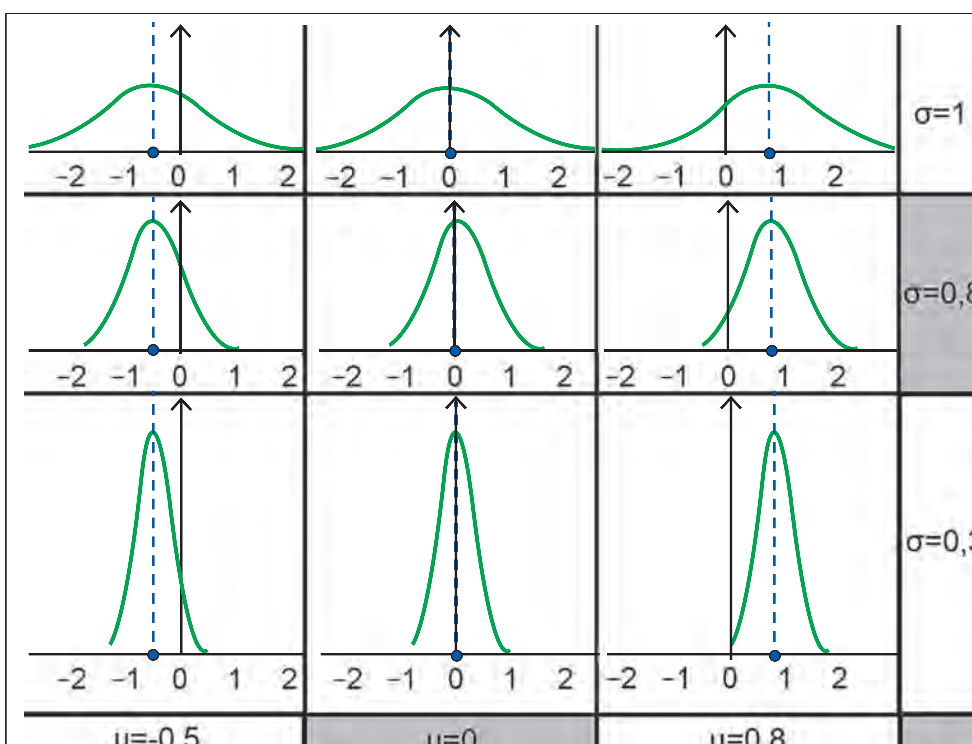
Όμοια μορφή έχουν και τα πολύγωνα σχετικών συχνοτήτων των παραπάνω παραδειγμάτων, σχεδιασμένα σε κατάλληλη κλίμακα (καθώς το 100% αντιστοιχεί στο 1).

Στη φύση συχνά συναντούμε κατανομές συχνοτήτων μεγάλων δειγμάτων, των οποίων τα ιστογράμματα και τα πολύγωνα σχετικών συχνοτήτων έχουν παρόμοια μορφή και ιδιότητες με τα δύο προηγούμενα παραδείγματα. Σε τέτοιες περιπτώσεις, όταν το πλήθος είναι αρκετά μεγάλο και το πλήθος των κλάσεων κατάλληλα μεγάλο, το πολύγωνο σχετικών συχνοτήτων τείνει να μοιάσει με τη γραφική παράσταση της συνάρτησης $y = \dfrac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$, για κάποιες τιμές των παραμέτρων μ και σ (η σ είναι θετική). Αυτή τη γραφική παράσταση την ονομάζουμε γκαουσιανή καμπύλη και φαίνεται στο παρακάτω σχήμα.




Η ευθεία x = μ είναι ο άξονας συμμετρίας της γραφικής παράστασης της συνάρτησης. Συνεπώς, μιλώντας κάπως πρακτικά, η τιμή του σ καθορίζει πόσο «απλωμένη» ή «μαζεμένη» είναι η καμπύλη γύρω από το μ και ποιο είναι το μέγιστο «ύψος» της. Η τιμή του μ επηρεάζει τη «θέση» της καμπύλης καθώς καθορίζει τη θέση του άξονα συμμετρίας της. Όταν το σ μεγαλώνει, το σχήμα της μοιάζει να είναι πιο «απλωμένο» πάνω από τον οριζόντιο άξονα και το μέγιστο «ύψος» της καμπύλης μικραίνει. Τα αντίστροφα συμβαίνουν όταν το σ μικραίνει. Ο ρόλος των τιμών των παραμέτρων μ και σ στη μορφή της γραφικής παράστασης φαίνεται και στα παρακάτω παραδείγματα.

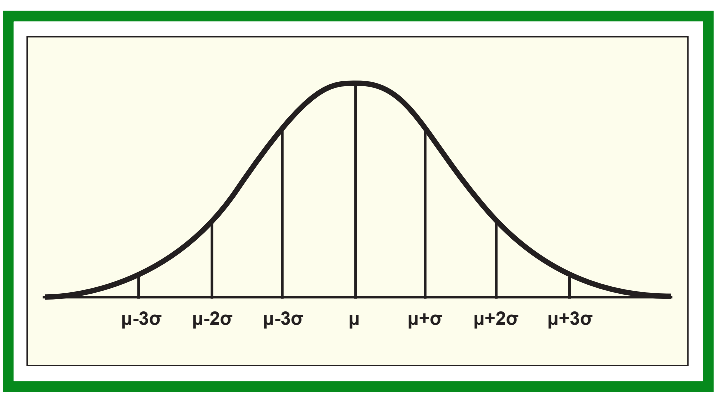
$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$




Η κανονική κατανομή είναι το μαθηματικό μοντέλο που χρησιμοποιούμε για να περιγράψουμε πώς κατανέμονται σε έναν ιδεατό, άπειρο πληθυσμό οι τιμές μεταβλητών, όπως εκείνες στα δύο παραδείγματά μας. Για την κανονική κατανομή, το μ εκφράζει το κέντρο συμμετρίας της και το σ είναι ένας δείκτης διασποράς της.

Έστω μία μεταβλητή που μοντελοποιείται από την κανονική κατανομή με μ και σ. Αν επιλέξουμε τυχαία ένα στοιχείο/άτομο από τον άπειρο πληθυσμό και κοιτάξουμε την τιμή της μεταβλητής για αυτό το στοιχείο/άτομο, τότε η πιθανότητα αυτή η τιμή να βρίσκεται σε ένα διάστημα (α,β) ισούται με το εμβαδόν μεταξύ της γκαουσιανής καμπύλης και του άξονα x΄x ανάμεσα στα α και β. Αποδεικνύεται ότι:

- η πιθανότητα του ενδεχομένου «η τιμή της μεταβλητής είναι μεγαλύτερη από μ» ισούται με 0,5 και, λόγω συμμετρίας, η πιθανότητα του ενδεχομένου «η τιμή της μεταβλητής είναι μικρότερη από μ» ισούται, επίσης με 0,5.
- η πιθανότητα του ενδεχομένου «η τιμή της μεταβλητής είναι στο διάστημα (μ - σ, μ + σ) » ισούται κατά προσέγγιση με 0,68 ή 68%.
- η πιθανότητα του ενδεχομένου «η τιμή της μεταβλητής είναι στο διάστημα (μ - 2σ , μ + 2σ) » ισούται κατά προσέγγιση με 0,95 ή 95%
- η πιθανότητα του ενδεχομένου «η τιμή της μεταβλητής είναι στο διάστημα (μ - 3σ , μ + 3σ)» ισούται κατά προσέγγιση με 0,997 ή 99,7%.



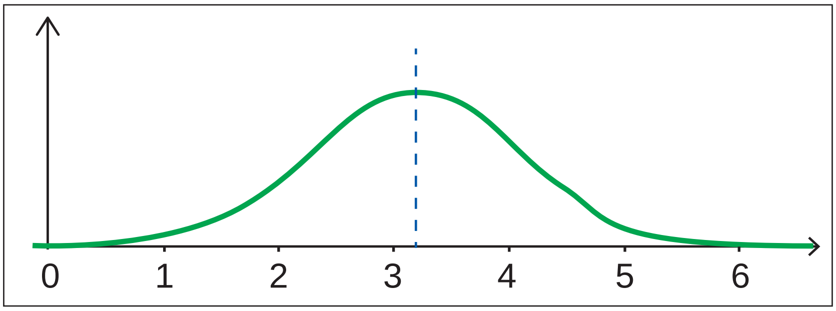

Από τα προηγούμενα και λόγω συμμετρίας της κανονικής κατανομής, υπολογίζουμε τις πιθανότητες και άλλων ενδεχομένων. Π.χ. η πιθανότητα του ενδεχομένου «η τιμή της μεταβλητής είναι στο διάστημα (μ - σ , μ)» ισούται κατά προσέγγιση με 0,68 : 2 = 0,34.

Λόγω στατιστικής ομαλότητας, σε ένα μεγάλο δείγμα (ή σε ολόκληρο τον υπαρκτό πληθυσμό) μπορούμε να εκτιμήσουμε ότι:

- το ποσοστό των στοιχείων/ατόμων για τα οποία η τιμή της μεταβλητής είναι στο διάστημα (μ - σ , μ + σ), είναι περίπου το 68% του πληθυσμού
- το ποσοστό των στοιχείων/ατόμων για τα οποία η τιμή της μεταβλητής είναι στο διάστημα (μ - 2σ , μ + 2σ), είναι περίπου το 95% του πληθυσμού
- το ποσοστό των στοιχείων/ατόμων για τα οποία η τιμή της μεταβλητής είναι στο διάστημα (μ - 3σ , μ + 3σ), είναι περίπου το 99,7% του πληθυσμού.

Λόγω συμμετρίας μπορούμε να κάνουμε και άλλες παρόμοιες εκτιμήσεις.

Επίσης, στην περίπτωση μεγάλου δείγματος (ή ολόκληρου του υπαρκτού πληθυσμού) η δειγματική μέση τιμή της μεταβλητής είναι κοντά στο μ και η δειγματική τυπική απόκλισή της είναι κοντά στο σ. Για παράδειγμα, στην περίπτωση του δείγματος των νεογέννητων, η μεταβλητή «βάρος νεογέννητου στις ΗΠΑ» μοντελοποιείται από την κανονική κατανομή με $\mu = 3{,}18$ και $\sigma = 0{,}91$ (βλέπε το παρακάτω σχήμα). Η μέση τιμή του βάρος των νεογέννητων του δείγματος είναι κοντά στο 3,18 και η τυπική απόκλισή του είναι κοντά στο 0,91. Το ίδιο ισχύει και για τη μέση τιμή και την τυπική απόκλιση του βάρους ολόκληρου του υπαρκτού πληθυσμού των νεογέννητων στις ΗΠΑ.




## Εφαρμογές

### Εφαρμογή 1

Ο πόρτες που κατασκευάζει μια εταιρία είναι τυποποιημένες με ύψος 183 cm. Θεωρούμε ότι τα ύψη των ενηλίκων στην Ελλάδα ακολουθούν κανονική κατανομή με μέση τιμή 171 cm και τυπική απόκλιση 6 cm:

1) Παίρνουμε ένα πολύ μεγάλο δείγμα ενηλίκων στην Ελλάδα. Να εκτιμήσετε το ποσοστό των ενήλικων του δείγματος που είναι ψηλότεροι από την πόρτα.

2) Να βρείτε ποιο πρέπει να είναι το ύψος της πόρτας, ώστε αν επιλέξουμε τυχαία έναν ενήλικα στην Ελλάδα, η πιθανότητα να είναι ψηλότερος/η από την πόρτα να είναι περίπου 0,15%;

**<u>Λύση</u>**

1) Τα ύψη των ενήλικων ανθρώπων ακολουθούν κανονική κατανομή με $\mu$=171cm και $\sigma$ = 6 cm. Ισχύουν:

$$\mu - 2\sigma = 171 - 2\cdot 6 = 159 \text{ και } \mu + 2\sigma = 171+2\cdot 6 = 183$$

Παίρνουμε ένα πολύ μεγάλο δείγμα ενηλίκων στην Ελλάδα. Θέλουμε να εκτιμήσουμε το ποσοστό των ατόμων του δείγματος που είναι ψηλότερα από την πόρτα, δηλαδή από 183 cm.

Το ποσοστό των ατόμων του δείγματος που το ύψος τους (σε cm) ανήκει στο διάστημα (159 , 183) εκτιμάται σε 95%.

Λόγω συμμετρίας της κανονικής κατανομής εκτιμούμε:

- Το ποσοστό των ατόμων του δείγματος που το ύψος τους ανήκει στο διάστημα (171 , 183) σε 95 : 2 = 47,5 %.
- Το ποσοστό των ατόμων του δείγματος που είναι ψηλότερα από 171 cm σε 50%.

Επομένως, το ποσοστό των ατόμων του δείγματος που είναι ψηλότερα από 183 cm, άρα και από την πόρτα, εκτιμάται σε 50% - 47,7% = 2,5%

2) Υποθέτουμε ότι η πόρτα έχει ύψος υ cm.

Έστω ένας τυχαίος ενήλικας και το ενδεχόμενο «ο τυχαίος ενήλικας έχει ύψος μεγαλύτερο από υ cm». Θέλουμε να βρούμε την τιμή του υ, ώστε η πιθανότητα του ενδεχομένου να είναι περίπου 0,15%.

Παρατηρούμε ότι 0,15% = 50% - 49,85% .

Επίσης ισχύει ότι $\mu+3\sigma = 171+3\cdot 6 = 189$.

Η πιθανότητα ο τυχαίος ενήλικας να είναι ψηλότερος από 171 cm είναι 50% και η πιθανότητα να έχει ύψος (σε εκατοστά) που ανήκει στο διάστημα $(\mu, \mu+3\sigma)=(171,189)$ είναι περίπου ίση με $\frac{99,7\%}{2} = 49,85\%$ , λόγω συμμετρίας της κανονικής κατανομής.

Άρα, η πιθανότητα ο τυχαίος ενήλικας να είναι ψηλότερος από 189 cm είναι περίπου 50% - 49,85% = 0,15%.

Άρα η πόρτα πρέπει να έχει ύψος υ = 189 cm.

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Το σύνολο των μαθητών/τριών μιας πόλης, ρωτήθηκαν για τον χρόνο που κάνουν να πάνε από το σπίτι στο σχολείο. Το 50% περίπου έδωσε απαντήσεις που ήταν από 12 λεπτά και πάνω, ενώ το 16% περίπου έδωσε απαντήσεις που ήταν από 10 λεπτά και κάτω. Υποθέτουμε ότι η κατανομή του χρόνου της διαδρομής σπίτι-σχολείο των μαθητών είναι κανονική.

1) Να εκτιμήσετε τον μέσο χρόνο διαδρομής σπίτι-σχολείο, των μαθητών και την τυπική απόκλιση του χρόνου διαδρομής τους, σύμφωνα με τις απαντήσεις τους.

2) Αν οι μαθητές/τριες της πόλης είναι 4.000, να εκτιμήσετε πόσοι/ες απάντησαν ότι έχουν χρόνο διαδρομής σπίτι-σχολείο μεταξύ 14 και 16 λεπτών;

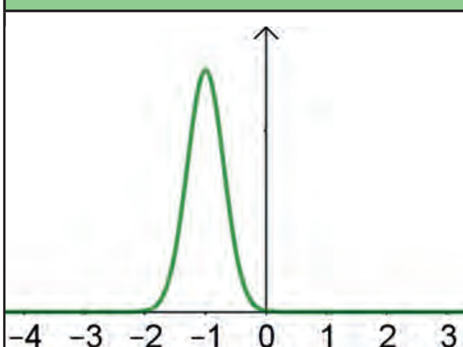
**2)** Υποθέτουμε ότι το βάρος των μαθητών λυκείου ακολουθεί κανονική κατανομή και παίρνουμε ένα μεγάλο δείγμα μαθητών λυκείου. Το 50% των μαθητών του δείγματος έχουν βάρος το πολύ 65 Kg, ενώ περίπου το 47,5% αυτών έχουν βάρος από 65 Kg έως 75 Kg.

1) Να εκτιμήσετε τη μέση τιμή και την τυπική απόκλιση του βάρους των μαθητών του δείγματος.

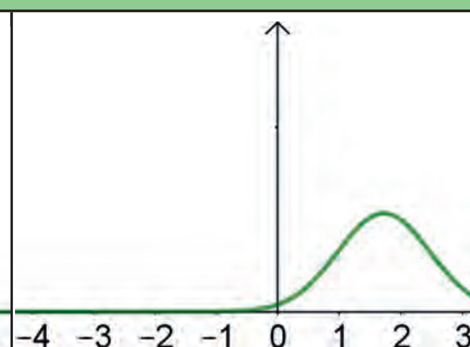
2) Να εκτιμήσετε το ποσοστό των μαθητών του δείγματος που έχουν βάρος από 55 Kg έως 70 Kg.

**3)** 1) Να αντιστοιχίσετε τις γραφικές παραστάσεις των συνάρτησεων που είναι μοντέλα κανονικών κατανομών και περιέχονται στον παρακάτω πίνακα με τα ζεύγη τιμών των παραμέτρων σ και μ που ακολουθούν. Ο κατακόρυφος άξονας των συστημάτων συντεταγμένων ακολουθεί την ίδια κλίμακα σε όλες τις περιπτώσεις.

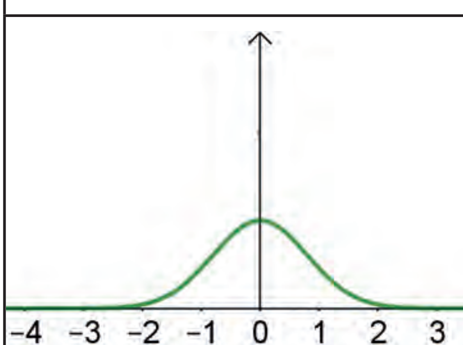
**Πίνακας**



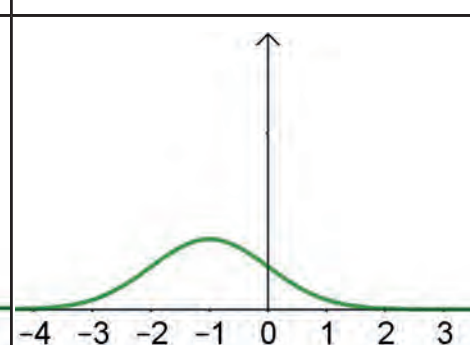

**(α)**




**(β)**




**(γ)**




**(δ)**

Α. μ = -1, σ = 1    Β. μ = -1, σ = 0,3    Γ. μ >0, σ = 0,75

Δ. μ = 0, σ < 1

2) Να συγκρίνετε την τιμή του σ στο σχήμα (γ) με το 0,3.

## Πρόσθετο Υλικό

1) Η κανονική κατανομή ανακαλύφθηκε το 1720 από τον μαθηματικό Abraham de Moivre, στην προσπάθειά του να λύσει προβλήματα παιγνίων τύχης. Εκατόν πενήντα χρόνια αργότερα περί το 1870 ο Adolph Quetelet, Βέλγος μαθηματικός, χρησιμοποιεί την καμπύλη της κανονικής κατανομής ως το ιδεώδες οριακό ιστόγραμμα - πρότυπο, προς το οποίο συγκρίνονται τα ιστογράμματα δεδομένων.

   Με τη χρήση κατάλληλου εκπαιδευτικού λογισμικού να σχεδιάσετε την καμπύλη (δηλ. τη γραφική παράσταση) της συνάρτησης $f(x) = \dfrac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$, για διάφορες τιμές των φυσικών αριθμών μ και σ και να κάνετε παρατηρήσεις για τη μορφή της.

# ΕΝΟΤΗΤΑ 2.5 ΠΙΝΑΚΕΣ ΣΥΝΑΦΕΙΑΣ ΚΑΙ ΡΑΒΔΟΓΡΑΜΜΑΤΑ

## Διερεύνηση

***Κατανομή μαθητών/μαθητριών σε ομάδες προσανατολισμού σπουδών***

Σε ένα Γενικό Λύκειο καταγράφηκαν το φύλο και η ομάδα προσανατολισμού σπουδών που επέλεξαν οι 50 μαθητές/μαθήτριες της Β' τάξης. Συγκεκριμένα, 11 αγόρια επέλεξαν την ομάδα προσανατολισμού Ανθρωπιστικών Σπουδών και 18 αγόρια επέλεξαν την ομάδα προσανατολισμού Θετικών Σπουδών. Αντίστοιχα, 9 κορίτσια επέλεξαν την ομάδα προσανατολισμού Ανθρωπιστικών Σπουδών και 12 κορίτσια την ομάδα προσανατολισμού Θετικών Σπουδών.

**α)** Να οργανώσετε τα παραπάνω δεδομένα σε έναν πίνακα συνάφειας (όπως δίνεται παρακάτω) που να περιέχει τις απόλυτες συχνότητες και τις σχετικές συχνότητες:

- ως προς το σύνολο των μαθητών/τριών της Β΄ Λυκείου
- ως προς το φύλο (γραμμές)
- ως προς την ομάδα προσανατολισμού Σπουδών (στήλες)

| | | **Ομάδα Προσανατολισμού Σπουδών** | | |
|---|---|---|---|---|
| | | **Ανθρωπιστικών** | **Θετικών** | **Σύνολο** |
| **Φύλο** | **Αγόρια** | | | |
| | **Κορίτσια** | | | |
| | **Σύνολο** | | | |

**β)** Να κατασκευάσετε κατάλληλα γραφήματα συχνοτήτων και σχετικών συχνοτήτων

- ως προς το σύνολο των μαθητών/τριών της Β Λυκείου
- ως προς το φύλο (γραμμές)
- ως προς την ομάδα προσανατολισμού Σπουδών (στήλες)

**γ)** Θα μπορούσατε να πείτε εάν το φύλο παίζει ρόλο στην επιλογή της ομάδας προσανατολισμού σπουδών;

## Βασικές μαθηματικές έννοιες – Ιδέες – Διεργασίες

Στα διάφορα προβλήματα που εξετάσαμε έως τώρα στη Στατιστική ασχοληθήκαμε κάθε φορά με μία μόνο μεταβλητή ξεχωριστά, ποιοτική ή ποσοτική, π.χ. με το φύλο των μαθητών, με τον χρόνο των αθλητών σε έναν αγώνα δρόμου, τον μηνιαίο μισθό των υπαλλήλων μιας εταιρείας. Μια πραγματική στατιστική ανάλυση δεν περιορίζεται συνήθως στη μελέτη μιας μεταβλητής, αλλά απαιτείται η μελέτη της σχέσης μεταξύ δύο ή περισσοτέρων μεταβλητών. Στην παράγραφο αυτή θα ασχοληθούμε με την αναζήτηση σχέσεων ανάμεσα σε δύο ποιοτικές μεταβλητές.

***πίνακας συνάφειας συχνοτήτων***

Ας δούμε τις βασικές ιδέες της διερεύνησης ενός προβλήματος με δύο ποιοτικές μεταβλητές μέσα από ένα παράδειγμα. Σε μια μελέτη που αφορά τη χρήση ζώνης ασφαλείας από οδηγούς ηλικίας 18 – 24 ετών μετρήθηκαν 198 οδηγοί σύμφωνα με το φύλο και το ενδεχόμενο να είχαν ή όχι κάποιο τροχαίο ατύχημα τα τελευταία πέντε χρόνια. Η εύρεση της πιθανής σχέσης μεταξύ του φύλου (χαρακτηριστικό Α) και τροχαίου ατυχήματος (χαρακτηριστικό Β) επιτυγχάνεται μέσω της κατασκευής του πίνακα συνάφειας (contingency table), ο οποίος είναι ένας πίνακας συχνοτήτων της μορφής:

| | | **Τροχαίο Ατύχημα** | | |
|---|---|---|---|---|
| | | **Ναι** | **Όχι** | **Σύνολο** |
| **Φύλο** | **Άνδρες** | **69** | **58** | **127** |
| | **Γυναίκες** | **27** | **44** | **71** |
| | **Σύνολο** | **96** | **102** | **198** |

*Πίνακας 1: Πίνακας Συνάφειας απόλυτων συχνοτήτων φύλου και συμμετοχής σε ατύχημα.*

Οι γραμμές του πίνακα συνάφειας (Πίνακας 1) αντιστοιχούν  στις κατηγορίες της μίας ποιοτικής μεταβλητής (π.χ. του φύλου) και οι στήλες στις κατηγορίες της άλλης ποιοτικής μεταβλητής (π.χ. του ατυχήματος), ενώ στο εσωτερικό του παρατίθενται οι συχνότητες που αντιστοιχούν σε όλους τους δυνατούς συνδυασμούς των κατηγοριών των δύο μεταβλητών.

***πίνακες συνάφειας σχετικών συχνοτήτων***

Το πλήθος των ατόμων στις κατηγορίες του Πίνακα 1 είναι διαφορετικό, οπότε δεν μπορούμε να εξηγήσουμε ποια είναι η σχέση μεταξύ φύλου και συμμετοχής σε ατύχημα. Προκειμένου να κατανοήσουμε καλύτερα αυτή τη σχέση,  θα πρέπει να εξετάσουμε τα ποσοστά αντί για τις συχνότητες. Ο πίνακας 1 μπορεί να μετατραπεί σε πίνακα σχετικών συχνοτήτων (%) ως προς το σύνολο των παρατηρήσεων του δείγματος ν (Πίνακας 2), ως προς το πλήθος των παρατηρήσεων του χαρακτηριστικού Α (Πίνακας 3) και ως προς το πλήθος των παρατηρήσεων του χαρακτηριστικού Β (Πίνακας 4). Από τον Πίνακα 2, προκύπτει ότι :


- Στο δείγμα της έρευνάς μας συμμετείχαν περισσότεροι άνδρες οδηγοί από γυναίκες οδηγούς.
- Οι περισσότεροι συμμετέχοντες στην έρευνα δε συμμετείχαν σε ατύχημα, ωστόσο δεν είναι και λίγοι εκείνοι που συμμετείχαν.

| | | Τροχαίο Ατύχημα | | |
|---|---|---|---|---|
| | | **Ναι** | **Όχι** | **Σύνολο** |
| **Φύλο** | **Άνδρες** | 34,9% | 29,3% | 64,2% |
| | **Γυναίκες** | 13,6% | 22,2% | 35,8% |
| | **Σύνολο** | 48,5% | 51,5% | 100% |

*Πίνακας 2: Πίνακας Συνάφειας σχετικών συχνοτήτων φύλου και ατυχήματος ως προς το σύνολο των παρατηρήσεων του δείγματος*

Η εμφάνιση των ποσοστών στο εσωτερικό ενός πίνακα συνάφειας μας δίνει τη δυνατότητα να διακρίνουμε τη μορφή της σχέσης που ενδεχομένως υπάρχει μεταξύ των μεταβλητών ενός πίνακα. Στην περίπτωση δύο μεταβλητών, η μια μεταβλητή μπορεί να παίξει τον ρόλο της ανεξάρτητης μεταβλητής (το φύλο) και η άλλη να παίξει τον ρόλο της εξαρτημένης (από το φύλο) μεταβλητής (το τροχαίο ατύχημα).

| | | Τροχαίο Ατύχημα | | |
|---|---|---|---|---|
| | | **Ναι** | **Όχι** | **Σύνολο** |
| **Φύλο** | **Άνδρες** | 54,3% | 45,7% | 100% |
| | **Γυναίκες** | 38% | 62% | 100% |
| | **Σύνολο** | | | |

*Πίνακας 3: Πίνακας Συνάφειας σχετικών συχνοτήτων φύλου και συμμετοχής σε ατύχημα ως προς το φύλο*

Από τον Πίνακα 3 φαίνεται ότι οι άνδρες οδηγοί τους δείγματος είχαν σε μεγαλύτερο ποσοστό ατυχήματα τα τελευταία πέντε χρόνια, σε σχέση με τις γυναίκες οδηγούς. Φαίνεται πως στο δείγμα μας υπάρχει μια σχέση ανάμεσα στο φύλο και στη συμμετοχή σε ατύχημα στο ιστορικό του οδηγού: οι άνδρες οδηγοί του δείγματος είναι πιο πιθανό να είχαν ατύχημα τα τελευταία 5 χρόνια. Το αν αυτή η σχέση που ανακαλύψαμε στο δείγμα μας θα μπορούσε να γενικευθεί στον πληθυσμό όλων των οδηγών δεν είναι προφανές. Εξαρτάται πρώτα απ' όλα από το αν το δείγμα που επιλέξαμε είναι αντιπροσωπευτικό του πληθυσμού. Ακόμα κι αν αυτό έχει εξασφαλιστεί με τη δειγματοληψία, το εύρημά μας θα μπορούσε να είναι τυχαίο (έτυχε ανάμεσα στους άνδρες οδηγούς που επιλέξαμε στο δείγμα να υπάρχουν πολλοί με ατύχημα τα τελευταία χρόνια σε σχέση με τον πληθυσμό των ανδρών ή κάτι αντίστοιχο για τις γυναίκες). Όσο μεγαλύτερο είναι το μέγεθος του δείγματος και όσο πιο μεγάλη είναι η διαφορά που βρήκαμε στο δείγμα μας, τόσο λιγότερο πιθανό είναι το εύρημα να είναι τυχαίο και με μεγαλύτερη εμπιστοσύνη μπορούμε να ισχυριστούμε ότι το εύρημα στο δείγμα μας αντανακλά μια διαφορά στον πληθυσμό.

| | | Τροχαίο Ατύχημα | | |
|---|---|---|---|---|
| | | **Ναι** | **Όχι** | **Σύνολο** |
| **Φύλο** | **Άνδρες** | **71,9%** | **56,9%** | |
| | **Γυναίκες** | **28,1%** | **43,1%** | |
| | **Σύνολο** | **100%** | **100%** | |

*Πίνακας 4: Πίνακας Συνάφειας σχετικών συχνοτήτων φύλου και της συμμετοχής σε ατύχημα ως προς τη συμμετοχή σε ατύχημα*

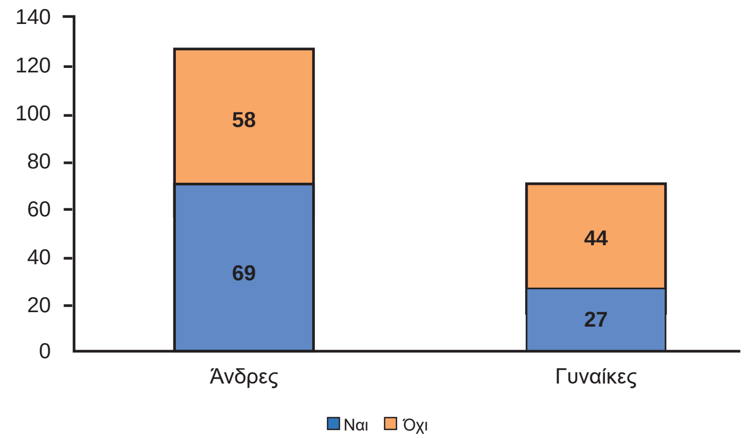
***Ραβδογράμματα***

Από τον πίνακα 4 φαίνεται ότι τόσο στο σύνολο των οδηγών που είχαν ατύχημα όσο και στο σύνολο των οδηγών που δεν είχαν ατύχημα, οι περισσότεροι οδηγοί ήταν άνδρες.

Όπως είδαμε στην περίπτωση της μίας ποιοτικής μεταβλητής οι πληροφορίες που αντλούνται από τους πίνακες συχνοτήτων ή σχετικών συχνοτήτων αναπαριστώνται με τα ραβδογράμματα και τα κυκλικά διαγράμματα. Στην περίπτωση των δύο ποιοτικών μεταβλητών, οπτικοποιούμε τα αποτελέσματα ενός πίνακα συνάφειας συχνοτήτων με το στοιβαγμένο ραβδόγραμμα (stacked bar plot) και τα αποτελέσματα ενός πίνακα συνάφειας σχετικών συχνοτήτων με το ομαδοποιημένο ραβδόγραμμα (clustered bar plot).

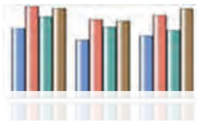
***Στοιβαγμένο Ραβδόγραμμα***

Το στοιβαγμένο ραβδόγραμμα συχνοτήτων (σχήμα 1) του πίνακα 2 αποτελείται από ορθογώνιες στήλες, οι βάσεις των οποίων βρίσκονται πάνω στον οριζόντιο άξονα, στον οποίο αντιστοιχούν οι κατηγορίες του ενός χαρακτηριστικού π.χ. του φύλου. Το πλήθος των ορθογώνιων στηλών είναι ίσο με το πλήθος των κατηγοριών του φύλου. Κάθε ορθογώνια στήλη διαιρείται σε τόσα τμήματα διαφορετικού χρώματος, όσες και οι κατηγορίες του άλλου χαρακτηριστικού (ατυχήματος). Συνήθως, στον οριζόντιο άξονα τοποθετούμε τις κατηγορίες της ανεξάρτητης μεταβλητής (φύλου, στο παράδειγμά μας) και στον κατακόρυφο άξονα τις κατηγορίες της εξαρτημένης μεταβλητής (ατυχήματος, στο παράδειγμά μας) ώστε να ελέγξουμε εάν το φύλο σχετίζεται με τα τροχαία ατυχήματα.

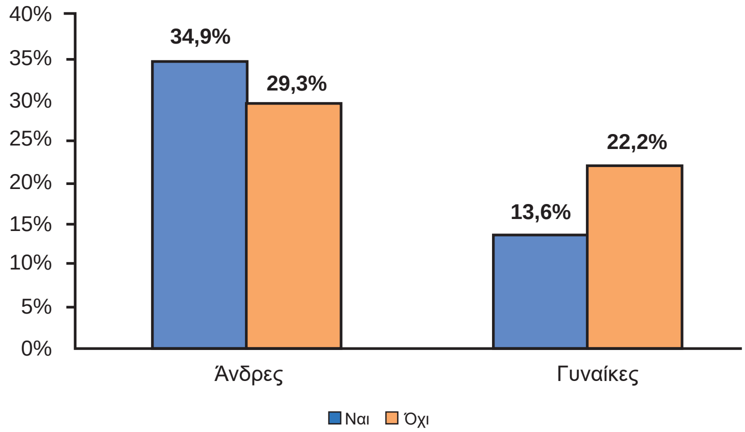



*Σχήμα 1: Στοιβαγμένο ραβδόγραμμα συχνοτήτων φύλου και συμμετοχής σε ατύχημα του Πίνακα 1*

***Ομαδοποιημένο Ραβδόγραμμα***



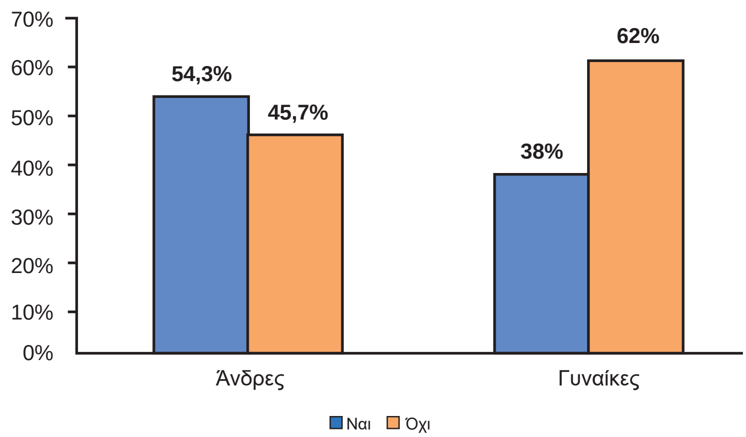
Το ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων αποτελείται από ορθογώνιες στήλες, οι βάσεις των οποίων βρίσκονται πάνω στον οριζόντιο άξονα στον οποίο αντιστοιχούν οι κατηγορίες του ενός χαρακτηριστικού π.χ. του φύλου. Σε κάθε κατηγορία του φύλου αντιστοιχούν τόσες ορθογώνιες στήλες διαφορετικού χρώματος όσες το πλήθος των κατηγοριών του άλλου χαρακτηριστικού π.χ. της συμμετοχής σε ατύχημα. Το ύψος κάθε στήλης είναι ίσο με την αντίστοιχη σχετική συχνότητα.




*Σχήμα 2: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων φύλου και ατυχήματος του πίνακα 2*

Στο σχήμα 2 φαίνονται όλα τα ποσοστά του πίνακα 2. Για παράδειγμα, εάν αθροίσετε όλα τα ποσοστά που φαίνονται στο σχήμα προκύπτει το 100%, ενώ εάν αθροίσετε τα ποσοστά των δύο ράβδων στους άνδρες οδηγούς προκύπτει το 64,2% που είναι το ποσοστό τους στο δείγμα. Επίσης, παρατηρούμε ότι οι ράβδοι σε κάθε κατηγορία του φύλου δεν έχουν το ίδιο μοτίβο. Δηλαδή, στους άνδρες οδηγούς οι περισσότεροι είχαν ατύχημα, ενώ στις γυναίκες οδηγούς οι περισσότερες δεν είχαν ατύχημα.

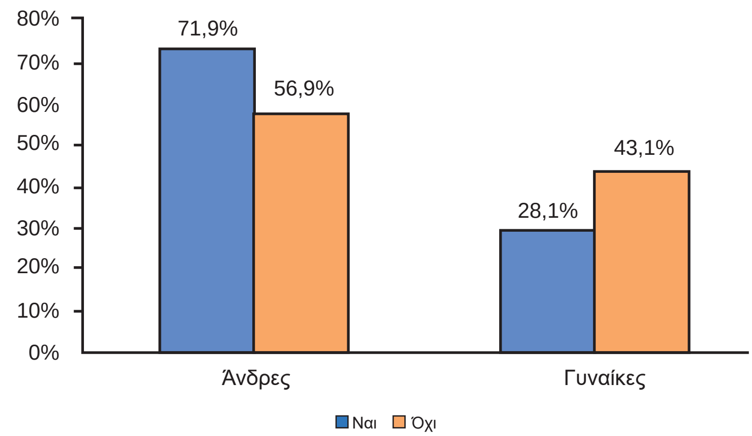
Στα σχήματα 3 και 4 φαίνονται τα ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων ως προς το φύλο και τη συμμετοχή σε ατύχημα, αντίστοιχα. Στο σχήμα 3 τα ποσοστά των ράβδων έχουν άθροισμα 100% τόσο ανάμεσα στους άνδρες όσο και ανάμεσα στις γυναίκες οδηγούς. Βλέπουμε, δηλαδή, στους άνδρες (γυναίκες) τα ποσοστά των οδηγών που είχαν/δεν είχαν ατύχημα στο σύνολο των ανδρών (γυναικών) οδηγών.




*Σχήμα 3: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων φύλου και της συμμετοχής σε ατύχημα ως προς το φύλο (αντ. Πίνακα 3)*

Από το σχήμα 3, το μοτίβο των ράβδων σε κάθε κατηγορία του φύλου είναι διαφορετικό, οπότε φαίνεται να υπάρχει μια σχέση ανάμεσα στο φύλο και στα ατυχήματα. Εάν το μοτίβο των ράβδων σε κάθε κατηγορία του φύλου ήταν περίπου ίδιο, αυτό θα ήταν μία ένδειξη ότι οι μεταβλητές φύλο και ατύχημα τα τελευταία πέντε χρόνια δε σχετίζονται.

Στο σχήμα 4 τα ποσοστά των ομοιόμορφα χρωματισμένων ράβδων έχουν άθροισμα 100%. Βλέπουμε, δηλαδή, τα ποσοστά των μπλε ράβδων να αντιστοιχούν στους άνδρες και γυναίκες οδηγούς που είχαν ατύχημα στο σύνολο των οδηγών που είχαν ατύχημα, ενώ τα ποσοστά των ροζ ράβδων αντιστοιχούν στους άνδρες και γυναίκες οδηγούς που δεν είχαν ατύχημα στο σύνολο των οδηγών που δεν είχαν ατύχημα. Από αυτό το σχήμα φαίνεται πιο εύκολα ότι το φύλο και η συμμετοχή σε ατύχημα σχετίζονται, αφού το ύψος των ράβδων σε κάθε κατηγορία του φύλου δεν είναι περίπου το ίδιο.




*Σχήμα 4: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων φύλου και συμμετοχής σε ατύχημα ως προς τη συμμετοχή σε ατύχημα (αντ. Πίνακα 4)*

***Σχέση Αιτιότητας***

Με ανάλογο τρόπο προκύπτουν στοιβαγμένα και ομαδοποιημένα ραβδογράμματα, εάν στον οριζόντιο άξονα τοποθετήσουμε τις κατηγορίες του ατυχήματος. Στο σημείο αυτό, καλό είναι να γίνει η παρακάτω διευκρίνιση. Το γεγονός ότι δύο μεταβλητές φαίνεται να σχετίζονται, όπως στο παράδειγμα μας, είναι λάθος να ερμηνεύεται με όρους αιτίου και αποτελέσματος, δηλαδή η ανεξάρτητη μεταβλητή να είναι η αιτία που προκλήθηκε η εξαρτημένη μεταβλητή (το αποτέλεσμα). Στο παράδειγμά μας δε θα ήταν σωστό για μια ασφαλιστική εταιρεία να ερμηνεύσει το αποτέλεσμα ως ότι το φύλο ενός οδηγού είναι υπεύθυνο για την πρόκληση ατυχημάτων. Με δεδομένο, λοιπόν, ότι δύο μεταβλητές A και B σχετίζονται, αυτό μπορεί να σημαίνει ότι:

- η μεταβλητή A είναι η αιτία για τη μεταβλητή B
- η μεταβλητή B είναι η αιτία για τη μεταβλητή A
- υπάρχει ένας τρίτος (συγχυτικός) παράγοντας ο οποίος να είναι η αιτία τόσο για το A, όσο και για το B.
- είναι απλά μια σύμπτωση, διότι απλά ένα τυχαίο γεγονός συνέβη στο δείγμα μας, ενώ στην πραγματικότητα, δηλαδή στον πληθυσμό, οι δύο μεταβλητές είναι ασυσχέτιστες.

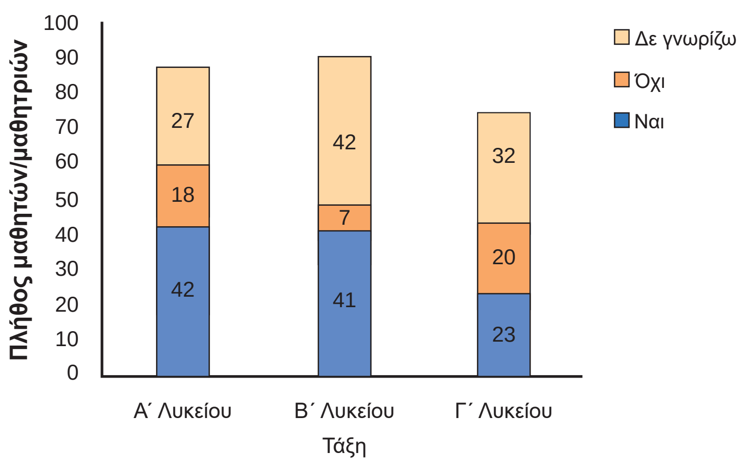
Για παράδειγμα, μια έρευνα έδειξε ότι τα μικρά παιδιά που κοιμούνται με αναμμένο το φως είναι πολύ πιθανότερο να εμφανίσουν μυωπία αργότερα. Μπορούμε άραγε να συμπεράνουμε ότι ο ύπνος με αναμμένο το φως προκαλεί μυωπία; Μεταγενέστερη έρευνα έδειξε ότι υπάρχει μια ισχυρή σχέση μεταξύ της γονικής μυωπίας και της εμφάνισης μυωπίας στα παιδιά τους, σημειώνοντας επίσης ότι οι μυωπικοί γονείς είναι πιο πιθανό να αφήσουν ένα φως στην κρεβατοκάμαρα των παιδιών τους. Ο συγχυτικός παράγοντας εδώ είναι η γονική μυωπία που επηρεάζει τόσο τη μεταβλητή «αναμμένο φως» όσο και τη μεταβλητή «μυωπία παιδιών».

Τέλος, αξίζει να αναφέρουμε το επόμενο παράδειγμα. Το αποτέλεσμα του τελευταίου τοπικού παιχνιδιού ποδοσφαίρου από την ομάδα των Washington Redskins πριν από τις προεδρικές εκλογές των Η.Π.Α. προέβλεπε το αποτέλεσμα της επόμενης προεδρικής εκλογής σε κάθε εκλογή από το 1940 έως το 2000. Όταν κέρδιζαν οι Redskins, το κόμμα του τότε προέδρου διατηρούσε την προεδρία, ενώ όταν οι Redskins έχαναν, κέρδιζε το κόμμα της αντιπολίτευσης. Φυσικά αυτό το εύρημα είναι μάλλον τυχαίο, αφού δεν είναι λογικό να υπάρχει σχέση ανάμεσα στα αποτελέσματα των τοπικών ποδοσφαιρικών αγώνων και στο αποτέλεσμα των προεδρικών εκλογών.

## Εφαρμογές

### Εφαρμογή 1

Στα πλαίσια μιας έρευνας που διενεργήθηκε σε 252 μαθητές/μαθήτριες ενός Λυκείου σχετικά με τους Παραολυμπιακούς αγώνες τέθηκε η ερώτηση: «Υπάρχει διαφορά μεταξύ των Special Olympics (SO) και των Παραολυμπιακών Αγώνων (ΠΑ);» Οι απαντήσεις των μαθητών/τριών του σχολείου αναπαριστώνται στο επόμενο στοιβαγμένο ραβδόγραμμα:




*Σχήμα 5: Στοιβαγμένο ραβδόγραμμα συχνοτήτων για το ερώτημα «Υπάρχει διαφορά μεταξύ των Special Olympics και των Παραολυμπιακών Αγώνων (ΠΑ);»*

**α)** Ποιες ήταν οι πιθανές απαντήσεις των μαθητών/τριών του Λυκείου;

**β)** Αθροίστε τις επιμέρους τιμές των μπλε ράβδων. Ποιο είναι το άθροισμά τους; Να ερμηνεύσετε το αποτέλεσμα. Να βρείτε τα ποσοστά των μαθητών/τριών του Λυκείου που έδωσαν κάθε διαφορετική απάντηση.

**γ)** Αθροίστε τις επιμέρους τιμές των ράβδων στη Β΄ τάξη. Ποιο είναι το άθροισμά τους; Να ερμηνεύσετε το αποτέλεσμα. Να εκφράσετε τις επιμέρους τιμές των ράβδων της Β΄ τάξης ως ποσοστά επί του συνόλου των μαθητών της Β΄ τάξης, ερμηνεύοντάς τα.

**δ)** Στο σύνολο των μαθητών/τριών του Λυκείου, πόσοι μαθητές/μαθήτριες από κάθε τάξη ήταν σίγουροι ότι δεν υπάρχει διαφορά μεταξύ (SO) και (ΠΑ); Τι ποσοστό ήταν από κάθε τάξη;

**ε)** Στο σύνολο των μαθητών/τριών του Λυκείου, ποιας τάξης οι μαθητές/μαθήτριες φαίνεται να είναι καλύτερα ενημερωμένοι ως προς το γεγονός ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ) και σε τι ποσοστό, επί του συνόλου των μαθητών/τριών του Λυκείου;

**Λύση**

**α)** Οι μαθητές/μαθήτριες απάντησαν «ναι», εάν πίστευαν ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ), «όχι», εάν πίστευαν ότι δεν υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ) και «δε γνωρίζω», εάν δε γνώριζαν την απάντηση.

**β)** Αθροίζοντας τις τιμές των μπλε ράβδων έχουμε 106 μαθητές/μαθήτριες (42 από την Α΄, 41 από τη Β΄ και 23 από τη Γ΄) που απάντησαν ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ). «Ναι» απάντησε το $\frac{106}{252} = 42,1\%$ , «όχι» το $\frac{45}{252} = 17,8\%$ και «δε γνωρίζω» το $\frac{101}{252} = 40,1\%$ των μαθητών/τριών.

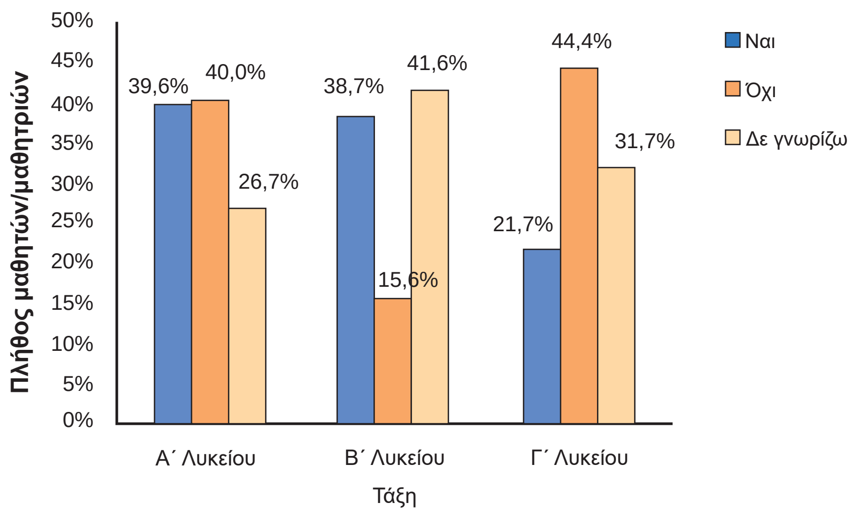
**γ)** Αθροίζοντας τις επιμέρους τιμές στη στήλη της Β΄ τάξης έχουμε 90 μαθητές/μαθήτριες της Β΄ Λυκείου εκ των οποίων οι 41 απάντησαν ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ), οι 7 απάντησαν ότι δεν υπάρχει διαφορά και οι 42 δε γνώριζαν την απάντηση. Τα ποσοστά των επιμέρους τιμών είναι: «Ναι» απάντησε περίπου το $\frac{41}{90} = 45,5\%$ , «όχι» απάντησε περίπου το $\frac{7}{90} = 7,8\%$ και «δε γνωρίζω» απάντησε περίπου το $\frac{42}{90} = 46,7\%$ των μαθητών/τριών της Β΄ Λυκείου.

**δ)** Από τους 252 μαθητές/μαθήτριες του Λυκείου, οι 18 από την Α΄ τάξη, οι 7 από τη Β΄ τάξη και οι 20 από τη Γ΄ τάξη απάντησαν «όχι». Τα ποσοστά είναι 7,1%, 2,8% και 7,9%, αντίστοιχα.

**ε)** Οι μαθητές/μαθήτριες που είναι καλύτερα ενημερωμένοι σύμφωνα με το γράφημα είναι οι 42 της Α΄ τάξης του Λυκείου. Το αντίστοιχο ποσοστό ως προς το σύνολο των μαθητών/μαθητριών του σχολείου είναι 16,7%. Ωστόσο και οι μαθητές/μαθήτριες της Β΄ τάξης θα μπορούσαμε να πούμε ότι είναι εξίσου καλά ενημερωμένοι.

## Εφαρμογή 2

Στα πλαίσια της ίδιας έρευνας για την ίδια ερώτηση κατασκευάζουμε το παρακάτω ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς τις απαντήσεις του ερωτήματος.




*Σχήμα 5: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς τις απαντήσεις του ερωτήματος «Υπάρχει διαφορά μεταξύ των Special Olympics και των Παραολυμπιακών Αγώνων (ΠΑ);»*

**α)** Από τους μαθητές/μαθήτριες του Λυκείου που απάντησαν Ναι στην ερώτηση, ποιο είναι το ποσοστό των μαθητών/τριών που αντιστοιχεί σε κάθε τάξη;

**β)** Από τους μαθητές/μαθήτριες του Λυκείου που απάντησαν Όχι στην ερώτηση, ποιο είναι το ποσοστό των μαθητών/τριών που αντιστοιχεί στην τάξη της Β Λυκείου και ποιο το ποσοστό που αντιστοιχεί στην τάξη της Γ΄ Λυκείου;

**γ)** Συγκρίνοντας τα ύψη των ράβδων του γραφήματος, μπορείτε να πείτε εάν τελικά οι απαντήσεις των μαθητών/τριών στο ερώτημα σχετίζονται με την τάξη στην οποία βρίσκονται;

### <u>Λύση</u>

**α)** Το ποσοστό των μαθητών/μαθητριών της Α΄ τάξης του Λυκείου που απάντησαν Ναι στην ερώτηση είναι 39,6%, ενώ για τις άλλες δύο τάξεις είναι 38,7% και 21,7%, αντίστοιχα.

**β)** Το ποσοστό των μαθητών/μαθητριών της Β΄ και Γ΄ τάξης του Λυκείου που απάντησαν Όχι στην ερώτηση είναι 15,6% και 44,4%, αντίστοιχα. Παρατηρούμε ότι το ποσοστό των μαθητών/μαθητριών της Γ΄ είναι πολύ μεγαλύτερο συγκριτικά με αυτό της Β΄. Αξιοσημείωτο είναι ότι οι μαθητές/μαθήτριες ήταν σίγουροι ότι δεν υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ), κάτι το οποίο δεν ισχύει.

**γ)** Από το σχήμα 5 βλέπουμε ότι οι ράβδοι σε κάθε τάξη δεν έχουν περίπου το ίδιο ύψος, γεγονός που μας υποδηλώνει ότι η τάξη και οι απαντήσεις των μαθητών σχετίζονται. Εάν όλες οι ράβδοι είχαν περίπου το ίδιο ύψος τότε ο παράγοντας της τάξης δε θα έπαιζε κανένα ρόλο στη διαμόρφωση των απαντήσεων των μαθητών/μαθητριών. Η καλύτερα ενημερωμένη τάξη είναι η Α΄ Λυκείου και η λιγότερο καλά ενημερωμένη είναι η Γ΄ τάξη. Αξίζει να σημειωθεί ότι οι μαθητές/μαθήτριες της Α΄ και της Γ΄ τάξης πίστευαν λανθασμένα ότι δεν υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ) σε πολύ μεγαλύτερο βαθμό από τη Β΄. Τέλος, να σημειώσουμε ότι είναι αρκετά μεγάλα τα ποσοστά της απάντησης «δε γνωρίζω» , γεγονός που υποδεικνύει την ελλιπή ενημέρωση των μαθητών/μαθητριών σχετικά με αυτές τις σημαντικές διοργανώσεις των Special Olympics και των Παραολυμπιακών Αγώνων.

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Για ποια από τα επόμενα ζεύγη μεταβλητών θα κατασκευάζατε πίνακα συνάφειας; Για εκείνα που θα επιλέξετε, να προσδιορίσετε την ανεξάρτητη και την εξαρτημένη μεταβλητή για την έρευνά σας.

**α)** φύλο και επίπεδο εκπαίδευσης

**β)** φύλο και χρήση διαδικτύου (σε ώρες/ημέρα)

**γ)** βαθμός στην Άλγεβρα και ώρες μελέτης (ανά εβδομάδα)

**δ)** ικανοποίηση από το σχολείο και φύλο

**ε)** ευτυχία στον γάμο και πίστη στα θαύματα

**2)** Σε μια έρευνα του τρόπου ζωής και των καθημερινών συνηθειών στη χώρα μας συλλέχθηκε ένα δείγμα 833 ατόμων απ' όλη την επικράτεια, προκειμένου να εξετάσουμε εάν υπάρχει σχέση μεταξύ του επίπεδου εκπαίδευσης (δευτεροβάθμια, αν έχουν ολοκληρώσει το Λύκειο και τριτοβάθμια, αν έχουν ολοκληρώσει τις σπουδές τους σε ΑΕΙ) και της συστηματικής σωματικής άσκησης τους (ναι: τουλάχιστον 2 φορές την εβδομάδα και όχι: το πολύ μια φορά την εβδομάδα) σε όλη τη διάρκεια ενός έτους. Τα δεδομένα οργανώθηκαν στον παρακάτω πίνακα συνάφειας:

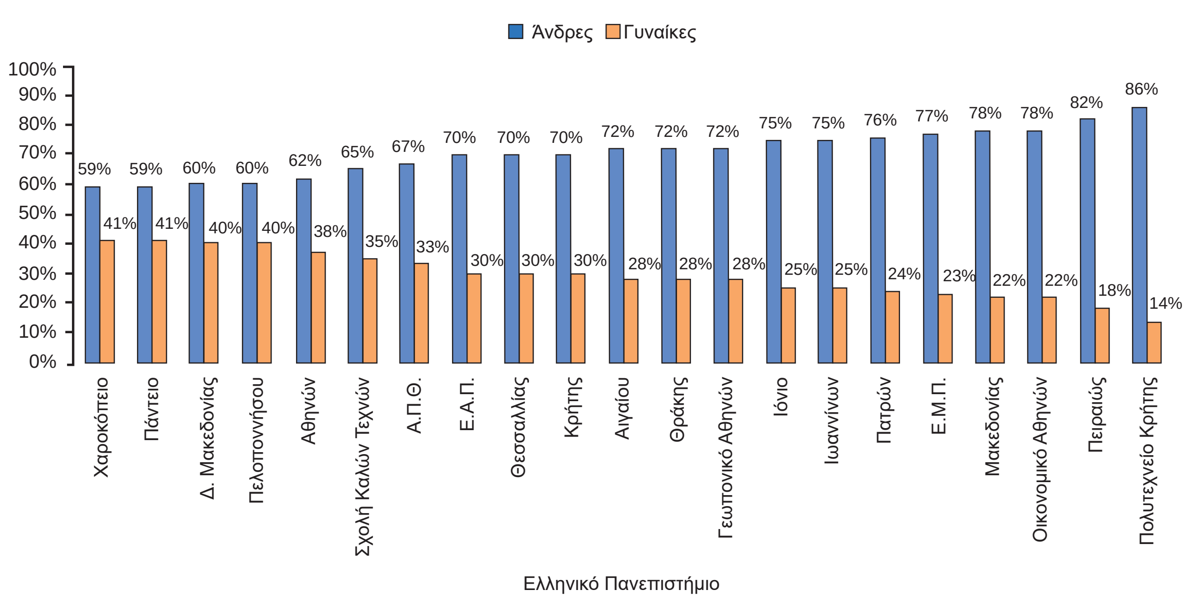
| | | Άσκηση | | |
|---|---|---|---|---|
| | | **Ναι** | **Όχι** | **Σύνολο** |
| **Εκπαίδευση** | **Δευτεροβάθμια** | **366** | **75** | |
| | **Τριτοβάθμια** | **303** | **89** | |
| | **Σύνολο** | | | |

**α)** Να συμπληρώσετε στον παραπάνω πίνακα τα κενά κελιά και να κατασκευάσετε το στοιβαγμένο ραβδόγραμμα των συχνοτήτων του πίνακα συνάφειας.

**β)** Να κατασκευάσετε τον πίνακα συνάφειας σχετικών συχνοτήτων ως προς το επίπεδο εκπαίδευσης και το αντίστοιχο ομαδοποιημένο ραβδόγραμμα του.

**γ)** Μπορείτε να υποστηρίξετε την άποψη ότι το επίπεδο εκπαίδευσης επηρεάζει το αν κάποιο άτομο ασκείται συστηματικά ή όχι;

**3)** Το επόμενο ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων δίνει τα ευρήματα μιας μελέτης που έγινε το 2016 στο Οικονομικό Πανεπιστήμιο Αθηνών για τη θέση των γυναικών στα Ελληνικά Πανεπιστήμια, η οποία δημοσιεύθηκε σε κυριακάτικη εφημερίδα. Πιο συγκεκριμένα, δίνονται τα ποσοστά ανδρών και γυναικών μόνιμων διδασκόντων στα Πανεπιστήμια.




**α)** Σε ποια Πανεπιστήμια παρατηρείται το μεγαλύτερο ποσοστό γυναικών και σε ποια το χαμηλότερο;

**β)** Τη συγκεκριμένη χρονιά δίδαξαν συνολικά 8.483 μόνιμοι διδάσκοντες εκ των οποίων οι 5.894 ήταν άνδρες και οι 2.589 ήταν γυναίκες. Να βρείτε σε ποια πανεπιστήμια το ποσοστό των γυναικών ξεπέρασε το γενικό μέσο ποσοστό τους;

**γ)** Ποιος πιστεύετε ότι μπορεί να ήταν ο τίτλος του άρθρου και για ποιον λόγο μια εφημερίδα το ανέδειξε ως ένα από τα κύρια κοινωνικά άρθρα της;

**4)** Στη Μυτιλήνη, ο δήμος εγκατέστησε ένα καινοτόμο πρόγραμμα ηλεκτρονικών υπηρεσιών του δήμου από το οποίο εξυπηρετούνται οι κάτοικοι του νησιού, ιδιοκτήτες και ενοικιαστές σπιτιών. Η ικανοποίηση των κατοίκων του νησιού από τις ηλεκτρονικές υπηρεσίες του δήμου φαίνονται από τα παρακάτω ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων (α) και (β).

**α)** Ποιο από τα δύο ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων αντιστοιχεί στο σύνολο των παρατηρήσεων του δείγματος; Τι δείχνει το άλλο ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων;

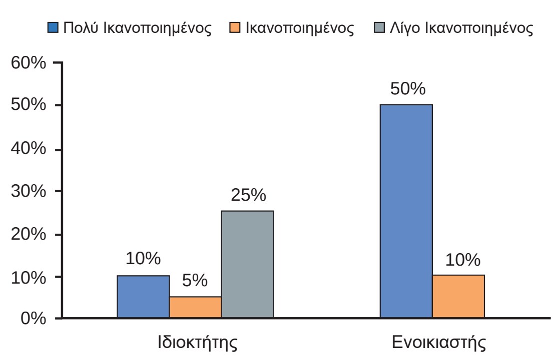
**β)** Ποιο ποσοστό του δείγματος είναι ιδιοκτήτες και ποιο ενοικιαστές σπιτιών;

**γ)** Ποιο ποσοστό του δείγματος είναι πολύ ικανοποιημένοι από τις ηλεκτρονικές υπηρεσίες του δήμου;

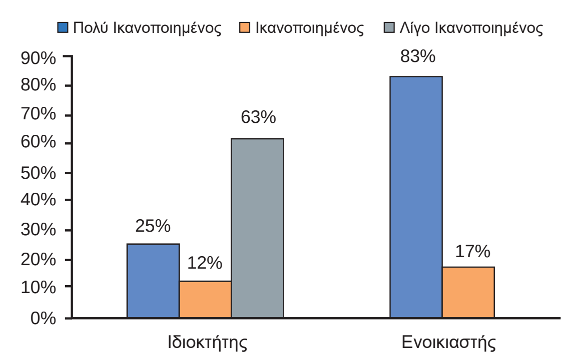
**δ)** Τι εκφράζουν τα ποσοστά 63% και 83% στο σχήμα (β);

**ε)** Αν το μέγεθος του δείγματος είναι 200 κάτοικοι, να βρείτε πόσοι είναι οι πολύ ικανοποιημένοι ιδιοκτήτες και πόσοι οι ικανοποιημένοι ενοικιαστές σπιτιών στο νησί.

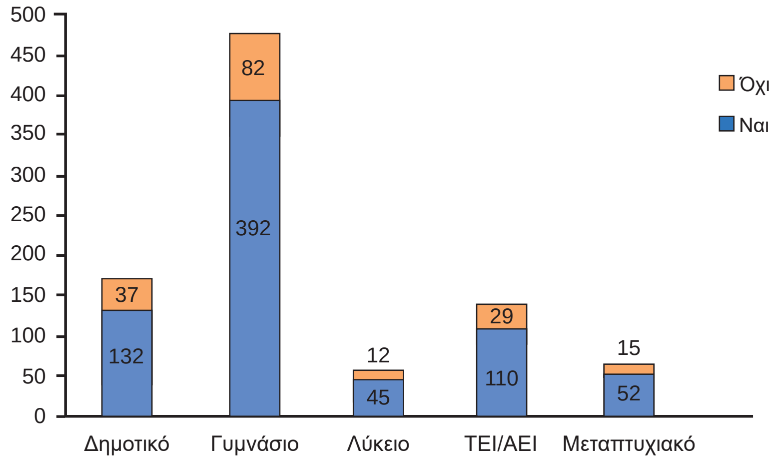
*Σχήμα (α)*




*Σχήμα (β)*




**5)** Πρόσφατα πραγματοποιήθηκε μια έρευνα προκειμένου να ερευνηθεί εάν υπάρχει σχέση μεταξύ της πίστης του ανθρώπου στα θαύματα και του επιπέδου εκπαίδευσης που έχει αποκτήσει. Τα αποτελέσματα της έρευνας συνοψίζονται στο επόμενο στοιβαγμένο ραβδόγραμμα συχνοτήτων.



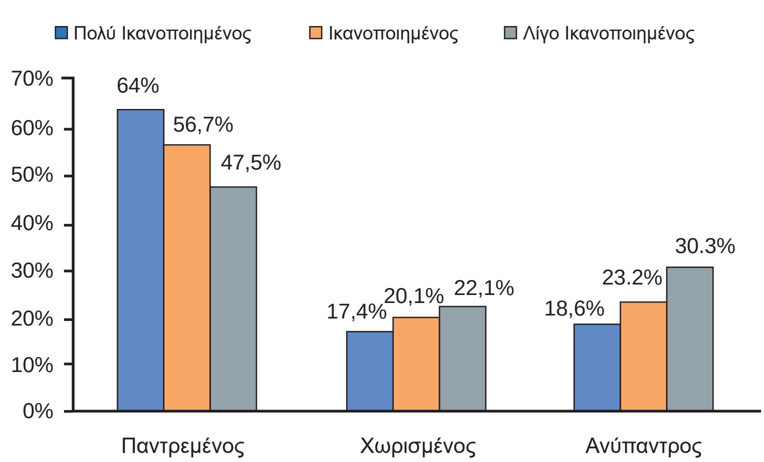

**α)** Τι ποσοστό αποφοίτων Γυμνασίου, Λυκείου και ΑΕΙ/ΤΕΙ έχουμε στο δείγμα μας;

**β)** Τι ποσοστό αποφοίτων Λυκείου φαίνεται να πιστεύουν στα θαύματα και τι ποσοστό αποφοίτων ΤΕΙ/ΑΕΙ φαίνεται ότι δεν πιστεύουν στα θαύματα;

**γ)** Να μετατρέψετε το παραπάνω στοιβαγμένο ραβδόγραμμα συχνοτήτων σε ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς την πίστη στα θαύματα.

**δ)** Με βάση τα αποτελέσματα της συγκεκριμένης έρευνας φαίνεται να υπάρχει κάποια σχέση μεταξύ του επιπέδου εκπαίδευσης και της πίστης στα θαύματα;

**6)** Μια κυριακάτικη εφημερίδα περιέχει ένα άρθρο σχετικά με την ικανοποίηση των ανθρώπων από την εργασία τους και την οικογενειακή τους κατάσταση μαζί με το παρακάτω γράφημα. Οι συντάκτες του άρθρου καταλήγουν στο συμπέρασμα ότι ο γάμος κάνει τους ανθρώπους περισσότερο ικανοποιημένους με την εργασία τους, δεδομένου ότι το 64% των πολύ ικανοποιημένων ανθρώπων είναι παντρεμένοι, ενώ μόνο το 18,6% δεν έχουν παντρευτεί ποτέ.




**α)** Συμφωνείτε με το συμπέρασμα του συντάκτη του άρθρου;

**β)** Συμφωνείτε με την αιτιολόγηση του συντάκτη του άρθρου;

# Πρόσθετο Υλικό

## Θέματα για διερεύνηση

**1)** Το φθινόπωρο του 1973 παρατηρήθηκε το ακόλουθο παράδοξο κατά την εισαγωγή μεταπτυχιακών φοιτητών στο Πανεπιστήμιο του Berkeley στην Καλιφόρνια. Τα στοιχεία της έρευνας έδειξαν ότι οι άνδρες που υπέβαλλαν αίτηση ήταν πιο πιθανό από ό,τι οι γυναίκες να γίνουν δεκτοί. Στον πίνακα (α) δίνονται τα στοιχεία αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές τη χρονιά του 1973. Στον πίνακα (β) δίνονται τα στοιχεία αιτούντων και εισακτέων ανδρών και γυναικών σε έξι διακεκριμένα τμήματα του Πανεπιστημίου του Berkeley για μεταπτυχιακές σπουδές τη χρονιά του 1973.

| | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό εισακτέων** |
|---|---|---|---|
| **Άνδρες** | 8442 | 3714 | …… |
| **Γυναίκες** | 4321 | 1512 | …... |

*Πίνακας (α): Πίνακας αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές στο Πανεπιστήμιο του Berkeley*

**α)** Συμπληρώστε τα ποσοστά των εισακτέων του πίνακα (α). Τι διαπιστώνετε, συγκρίνοντας τα ποσοστά των ανδρών και των γυναικών εισακτέων;

| | **Άνδρες** | | | **Γυναίκες** | | |
|---|---|---|---|---|---|---|
| | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό εισακτέων** | **Αιτούντες** | **Εισακτέοι** | **Ποσοστό εισακτέων** |
| **Τμήμα Α** | 825 | 512 | …….. | 108 | 89 | …….. |
| **Τμήμα Β** | 560 | 353 | …….. | 25 | 17 | …….. |
| **Τμήμα Γ** | 325 | 120 | …….. | 593 | 202 | …….. |
| **Τμήμα Δ** | 417 | 138 | …….. | 375 | 131 | …….. |
| **Τμήμα Ε** | 191 | 54 | …….. | 393 | 94 | …….. |
| **Τμήμα Ζ** | 373 | 23 | …….. | 341 | 24 | …….. |

*Πίνακας (β): Πίνακας αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές σε έξι διακεκριμένα τμήματα Πανεπιστήμιο του Berkeley*

**β)** Συμπληρώστε τα ποσοστά των εισακτέων του πίνακα (β) και συγκρίνετε:

**(ι)** τα ποσοστά των ανδρών εισακτέων στα τμήματα Γ και Ε με τα αντίστοιχα ποσοστά των γυναικών,

**(ιι)** τα ποσοστά των εισακτέων γυναικών στα τμήματα Α, Β, Δ και Ζ με τα αντίστοιχα ποσοστά των ανδρών.

**γ)** Τι διαπιστώνετε, συγκρίνοντας τα παραπάνω ποσοστά σε κάθε τμήμα του Πανεπιστήμιου με τα ποσοστά του ερωτήματος α; Υπάρχει κάποια σύγχυση;

**δ)** Σε τι είδους τμήματα ως προς την ανταγωνιστικότητα και το ποσοστό των εισακτέων φαίνεται να απευθύνονται οι γυναίκες και σε τι είδους τμήματα οι άνδρες;

**4)** Το σχολικό έτος 2017-2018 η ΕΛ. ΣΤΑΤ. (Ελληνική Στατιστική Αρχή) διοργάνωσε για πρώτη φορά έναν πανελλήνιο διαγωνισμό Στατιστικής. Στη Β΄ φάση οι διαγωνιζόμενοι κλήθηκαν να επεξεργαστούν τα δεδομένα μιας στατιστικής έρευνας χρήσης τεχνολογιών πληροφορικής και επικοινωνίας από 3351 πολίτες το πρώτο τρίμηνο του 2017. Ένα από τα αντικείμενα μελέτης ήταν η χρήση υπηρεσιών διακυβέρνησης από τους πολίτες, ανάλογα με την ηλικία τους, το επίπεδο εκπαίδευσης, τη γεωγραφική περιοχή και την κύρια ασχολία τους. Πιο συγκεκριμένα, οι πολίτες ρωτήθηκαν εάν λαμβάνουν πληροφορίες από ιστοσελίδες δημόσιων υπηρεσιών. Οι πιθανές απαντήσεις ήταν δύο: Ναι, εάν έχουν λάβει πληροφορίες από ιστοσελίδες δημόσιων υπηρεσιών και Όχι, εάν δεν έχουν κάνει τη συγκεκριμένη ενέργεια. Τα αποτελέσματα συνοψίζονται στους επόμενους πίνακες συνάφειας.

| | **Ναι** | **Όχι** | **Σύνολο** |
|---|---|---|---|
| **Άνδρας** | 1094 | 573 | 1667 |
| **Γυναίκα** | 1056 | 628 | 1684 |
| **Σύνολο** | 2150 | 1201 | 3351 |

*Πίνακας 1: Πίνακας Συνάφειας συχνοτήτων φύλου και χρήσης ηλ. Διακυβέρνησης*

| | **Ναι** | **Όχι** | **Σύνολο** |
|---|---|---|---|
| **16-34** | 616 | 293 | 909 |
| **35-54** | 1109 | 529 | 1638 |
| **>54** | 425 | 379 | 804 |
| **Σύνολο** | 2150 | 1201 | 3351 |

*Πίνακας 2: Πίνακας Συνάφειας συχνοτήτων ηλικίας και χρήσης ηλ. Διακυβέρνησης*

| | Ναι | Όχι | Σύνολο |
|---|---|---|---|
| **Έως Γυμνάσιο** | 175 | 360 | 535 |
| **Έως Λύκειο** | 829 | 622 | 1451 |
| **Έως Ανώτατη Εκπαίδευση** | 1146 | 219 | 1365 |
| **Σύνολο** | 2150 | 1201 | 3351 |

*Πίνακας 3: Πίνακας Συνάφειας συχνοτήτων επιπέδου εκπαίδευσης και χρήσης ηλ. Διακυβέρνησης*

[1]**Νήσοι Αιγαίου – Κρήτη**: Βόρειο Αιγαίο, Νότιο Αιγαίο και Κρήτη

[2]**Βόρεια Ελλάδα:** Ανατολική Μακεδονία και Θράκη, Κεντρική Μακεδονία, Δυτική Μακεδονία και Ήπειρος

[3]**Κεντρική Ελλάδα**: Θεσσαλία, Νησιά Ιονίου, Δυτική Ελλάδα, Λοιπή Στερεά Ελλάδα και Πελοπόννησος

| | Ναι | Όχι | Σύνολο |
|---|---|---|---|
| **Αττική** | 1002 | 454 | 1456 |
| **Νήσοι Αιγαίου – Κρήτη**[1] | 185 | 167 | 352 |
| **Βόρεια Ελλάδα**[2] | 556 | 313 | 869 |
| **Κεντρική Ελλάδα**[3] | 407 | 267 | 674 |
| **Σύνολο** | 2150 | 1201 | 3351 |

*Πίνακας 4: Πίνακας Συνάφειας συχνοτήτων γεωγραφικής περιοχής και χρήσης ηλ. Διακυβέρνησης*

| | Ναι | Όχι | Σύνολο |
|---|---|---|---|
| **Μισθωτός - Ελ.Επ.** | 1329 | 499 | 1828 |
| **Άνεργος** | 309 | 199 | 508 |
| **Μαθητής/Φοιτητής** | 165 | 105 | 270 |
| **Άλλο**[4] | 347 | 398 | 745 |
| **Σύνολο** | 2150 | 1201 | 3351 |

*Πίνακας 5: Πίνακας Συνάφειας συχνοτήτων γεωγραφικής περιοχής και χρήσης ηλ. Διακυβέρνησης ( [4]**Άλλο:** Συνταξιούχος, Νοικοκυρά, Στρατιώτης)*

Να γράψετε μια ερευνητική έκθεση με τις πληροφορίες που μπορείτε να αντλήσετε από τους παραπάνω πίνακες συνάφειας και την επεξεργασία τους (κατασκευή πινάκων συνάφειας σχετικών συχνοτήτων, κατασκευή κατάλληλων γραφημάτων), απαντώντας στα επόμενα ερωτήματα:

**α)** Ποιες είναι οι μεταβλητές της έρευνας και ποιο το είδος τους;

**β)** Να κατασκευάσετε κατάλληλο γράφημα για κάθε μεταβλητή χωριστά και να ερμηνεύσετε τα αποτελέσματα.

**γ)** Με κατάλληλα γραφήματα, να περιγράψετε τη σχέση της χρήσης υπηρεσιών ηλεκτρονικής διακυβέρνησης, ανάλογα με το φύλο, την ηλικία, το επίπεδο εκπαίδευσης, τη γεωγραφική περιοχή και την κύρια ασχολία των ερωτηθέντων.

**δ)** Μπορείτε να πείτε εάν η χρήση υπηρεσιών ηλεκτρονικής διακυβέρνησης συσχετίζεται με κάποιους από τους παράγοντες, όπως το φύλο, την ηλικία, το επίπεδο εκπαίδευσης, τη γεωγραφική περιοχή και την κύρια ασχολία.

**3)** Αποφασίζετε με τους φίλους σου να κάνετε μια έρευνα για ένα θέμα που σας απασχολεί ή συζητήσατε σε κάποιο άλλο μάθημα και θέλετε να δείτε τι γνώμη έχουν οι συμμαθητές σας. Συντάξτε ένα ερωτηματολόγιο με ερωτήσεις που αφορούν τα βασικά στοιχεία των συμμαθητών σας (φύλο, τάξη, ομάδα προσανατολισμού) και ερωτήσεις που αφορούν το θέμα που επιλέξατε. Φροντίστε οι ερωτήσεις να είναι σαφώς διατυπωμένες και οι πιθανές απαντήσεις να μην επικαλύπτονται. Χρησιμοποιώντας αυτά που μάθατε στη συγκεκριμένη παράγραφο, γράψτε μια σύντομη ερευνητική έκθεση παρουσιάζοντας τα αποτελέσματά σας. Για παράδειγμα, ο εθισμός στο διαδίκτυο είναι μια νέα μορφή εξάρτησης και αναφέρεται στην καταναγκαστική και υπερβολική χρήση του διαδικτύου και τον εκνευρισμό που προκαλείται από τη στέρηση της.

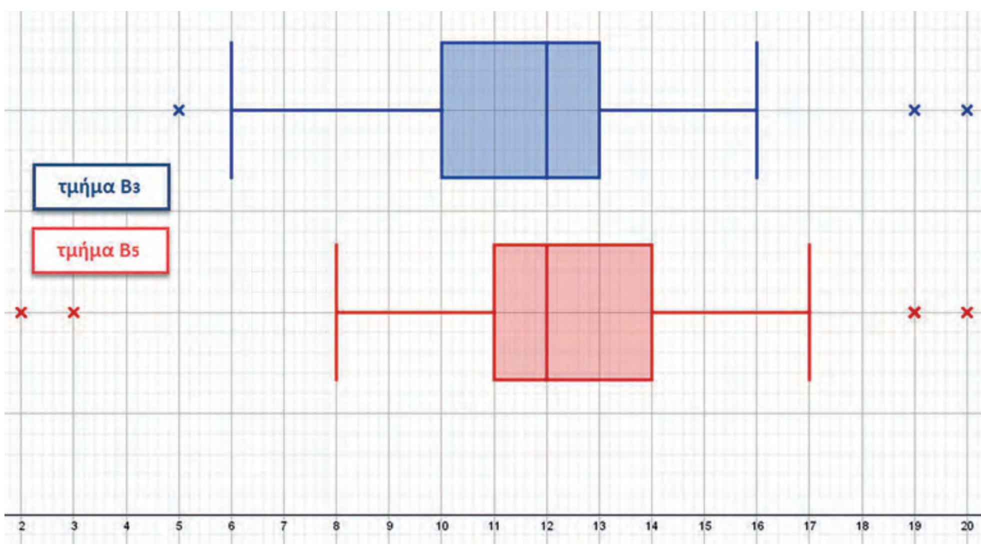
Αναζητήστε συσχετίσεις ανάμεσα στον χρόνο χρήσης του διαδικτύου (καθόλου, μέχρι μία ώρα την ημέρα, από 1 έως και 3 ώρες την ημέρα, πάνω από 3 ώρες την ημέρα) με διάφορους παράγοντες που πιστεύετε ότι επηρεάζουν την ενασχόληση με το διαδίκτυο.

# ΕΝΟΤΗΤΑ 2.6 ΣΥΓΚΡΙΣΗ ΠΟΣΟΤΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΣΤΙΣ ΚΑΤΗΓΟΡΙΕΣ ΕΝΟΣ ΠΟΙΟΤΙΚΟΥ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟΥ

## Διερεύνηση

***Το καλύτερο τμήμα***

Δύο τμήματα, το $B_3$ και το $B_5$ της Β΄ Λυκείου ενός σχολείου συμμετείχαν σε έναν μαθηματικό διαγωνισμό και οι βαθμοί τους παρουσιάζονται στα επόμενα θηκογράμματα.




1) Να βρείτε ποιο από τα δύο τμήματα έχει το μεγαλύτερο εύρος βαθμών;
2) Να βρείτε ποιο από τα τμήματα έχει το μεγαλύτερο ενδοτεταρτημοριακό εύρος βαθμών;
3) Ποιο τμήμα πιστεύετε ότι είναι το καλύτερο από τα δύο;

## Βασικές μαθηματικές έννοιες – Ιδέες – Διεργασίες

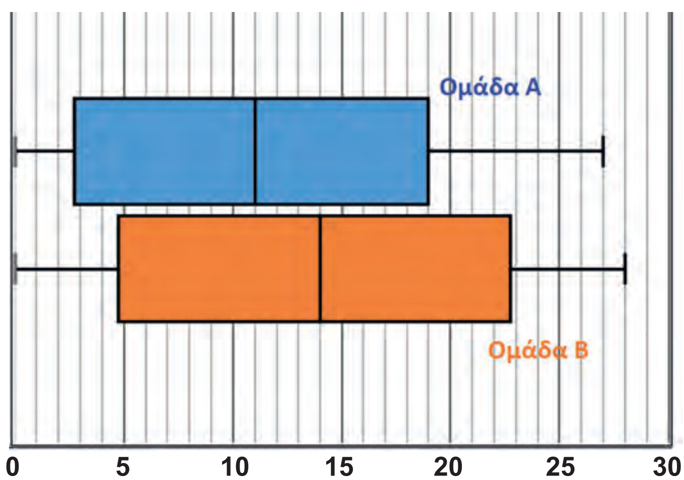
Σε προηγούμενη παράγραφο είδαμε αναλυτικά τα μέτρα θέσης και μεταβλητότητας για μια μεταβλητή. Εδώ, θα τα χρησιμοποιήσουμε για να συγκρίνουμε ομάδες δεδομένων.

Για παράδειγμα, παρακάτω βλέπουμε τα γκολ που έχουν σημειώσει οι παίκτες δύο ομάδων χάντμπωλ σε δέκα αγώνες. Η ομάδα Α έχει 20 παίκτες, ενώ η ομάδα Β έχει 18.

Ομάδα Α: 0, 0, 0, 1, 2, 5, 7, 8, 10, 11, 11, 17, 18, 18, 19, 19, 24, 26, 27, 27.

Ομάδα Β: 0, 0, 1, 4, 5, 7, 8, 10, 13, 15, 17, 20, 21, 22, 25, 26, 27, 28.

Οι δύο ομάδες έχουν σημειώσει συνολικά περίπου τα ίδια γκολ (250 η ομάδα Α και 249 η ομάδα Β). Επίσης, ο μέσος όρος των γκολ της ομάδας Α είναι 12,5 και της ομάδας Β είναι 13,83, περίπου. Η διαφορά των μέσων όρων προκύπτει από το γεγονός ότι η ομάδα Β, αν και έχει δύο παίκτες λιγότερους, έχει πετύχει σχεδόν τα ίδια γκολ. Μπορούμε, όμως με μεγαλύτερη λεπτομέρεια να δούμε, πώς η ομάδα Β το καταφέρνει αυτό; Στην προσπάθειά μας να συγκρίνουμε τις δύο ομάδες με κριτήριο τα γκολ που πέτυχε κάθε παίκτης, το παρακάτω θηκόγραμμα δίνει μία αρκετά σαφή εικόνα. Από τη θέση των δύο ορθογωνίων φαίνεται ότι οι «μεσαίοι σκόρερ» της ομάδας Β (δηλαδή εκείνοι που πετυχαίνουν έναν ούτε μεγάλο ούτε μικρό αριθμό γκολ») έχουν πετύχει περισσότερα γκολ από τους «μεσαίους σκόρερ» της ομάδας Α. Αυτό, ενδεχομένως, εξηγεί και γιατί η ομάδα Β με δύο λιγότερους παίκτες έχει πετύχει σχεδόν τα ίδια γκολ. Επίσης, συγκρίνοντας τα μήκη των ορθογωνίων βλέπουμε ότι για τους «μεσαίους σκόρερ» υπάρχει μεγαλύτερη μεταβολή των γκολ στην ομάδα Β (μεγαλύτερο μήκος), ενώ από τα μήκη των θηκογραμμάτων φαίνεται ότι το εύρος των γκολ της ομάδας Β είναι κατά 1 γκολ μεγαλύτερο (κάτι που οφείλεται στον παίκτη που σημείωσε 28 γκολ).




Σε ένα άλλο παράδειγμα, ο οργανισμός συγκοινωνιών σε ένα νησί του Αιγαίου μετρά τη διάρκεια ενός συγκεκριμένου δρομολογίου σε λεπτά (min) καθημερινά για τους μήνες Μάιο και Ιούλιο του ίδιου έτους, προκειμένου να συγκρίνει το επίπεδο υπηρεσιών του μεταξύ Μαΐου και Ιουλίου. Τα δειγματικά στατιστικά μέτρα για κάθε μήνα φαίνονται στον επόμενο πίνακα 1.

| | Μάιος | Ιούλιος |
|---|---|---|
| $\overline{x}$ | 106,67 min | 107,33 min |
| $\delta(Q_2)$ | 101 min | 106 min |
| $Q_1$ | 95 min | 95 min |
| $Q_3$ | 108 min | 118 min |
| $x_{min}$ | 81 min | 85 min |
| $x_{max}$ | 167 min | 155 min |
| S | 23,57 min | 16,15 min |

*Πίνακας 1: Δειγματικά στατιστικά μέτρα της διάρκειας ενός συγκεκριμένου δρομολογίου για τους μήνες Μάιο και Ιούλιο*

***Σύγκριση μέτρων θέσης της διάρκειας ενός συγκεκριμένου δρομολογίου για τους μήνες Μάιο και Ιούλιο***

Η δειγματική μέση τιμή της διάρκειας του συγκεκριμένου δρομολογίου του Μαΐου είναι 106,67 λεπτά, ενώ η αντίστοιχη δειγματική μέση τιμή για τον Ιούλιο είναι 107,33 λεπτά. Εκ πρώτης όψεως θα μπορούσε κάποιος να πει ότι συγκρίνοντας τις δύο δειγματικές μέσες τιμές η μέση διάρκεια του συγκεκριμένου δρομολογίου δεν άλλαξε αισθητά, οπότε η εταιρεία δε χρειάζεται να προχωρήσει σε κάποια βελτιωτική ενέργεια.

Η δειγματική διάμεσος για τον μήνα Μάιο είναι 101 λεπτά, ενώ για τον Ιούλιο η αντίστοιχη τιμή είναι 106 λεπτά. Αυτό σημαίνει ότι τον Μάιο τα μισά περίπου συγκεκριμένα δρομολόγια εκτελούνται σε χρόνο το πολύ 101 λεπτά, ενώ τα υπόλοιπα σε χρόνο τουλάχιστον 101 λεπτών. Αντίστοιχα, τον Ιούλιο τα μισά περίπου δρομολόγια εκτελούνται το πολύ σε 106 λεπτά, ενώ τα υπόλοιπα σε 106 λεπτά και περισσότερο.

***Σύγκριση μέτρων μεταβλητότητας της διάρκειας ενός συγκεκριμένου δρομολογίου για τους μήνες Μάιο και Ιούλιο***

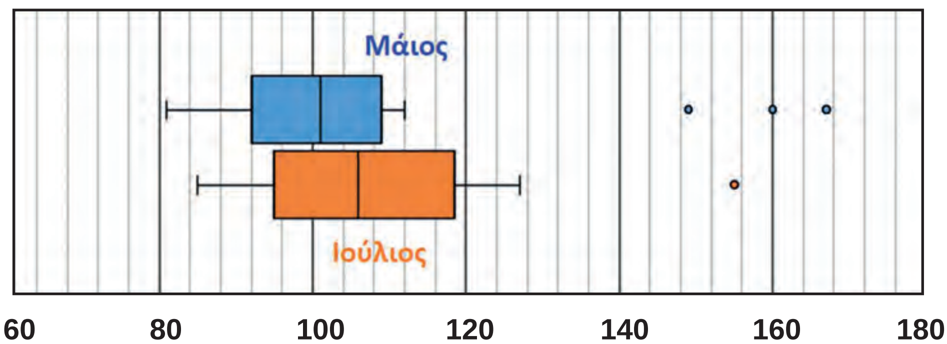
Το εύρος της διάρκειας του δρομολογίου τον Μάιο είναι 86 λεπτά και τον Ιούλιο είναι 70 λεπτά, κάτι που σημαίνει ότι η διαφορά της διάρκειας του συντομότερου με το μακρύτερο δρομολόγιο τον Μάιο είναι μεγαλύτερη. Από την άλλη μεριά, το ενδοτεταρτημοριακό εύρος της διάρκειας του δρομολογίου για το Μάιο είναι 13 λεπτά, ενώ για τον Ιούλιο είναι 23 λεπτά, κάτι που δηλώνει ότι το διάστημα $[Q_1, Q_3]$ είναι κατά 10 λεπτά μεγαλύτερο, από το αντίστοιχο του Μαΐου. Επομένως, παρατηρείται πιο σημαντική μεταβολή για τη διάρκεια των δρομολογίων του Ιουλίου, που βρίσκονται στο διάστημα $[Q_1, Q_3]$, δηλαδή για τα δρομολόγια, τα οποία μπορούν να χαρακτηριστούν ως «μέσης διάρκειας» δρομολόγια, σε σύγκριση με τα αντίστοιχα δρομολόγια του Μαΐου.

Αν δει κανείς, επιπλέον, τις τιμές των δύο δειγματικών τυπικών αποκλίσεων (23,57 λεπτά και 16,15 λεπτά), φαίνεται ότι διάρκειες των συνολικών δρομολογίων του Μαΐου ήταν πιο «διάσπαρτες» γύρω από τη μέση διάρκεια. Τέλος, αν υπολογίσουμε τους συντελεστές μεταβλητότητας για κάθε μήνα, τότε προκύπτουν $CV_M = 22{,}1\%$ περίπου και $CV_I = 15\%$, περίπου. Άρα, φαίνεται ότι η διάρκεια των δρομολογίων του Μαΐου, συνολικά, παρουσιάζει μεγαλύτερη μεταβλητότητα και μικρότερη ομοιογένεια από αυτά του Ιουλίου.

***Σύγκριση θηκογραμμάτων***

Βλέπουμε ότι συμπεράσματα που βγαίνουν από τη μελέτη των δειγμάτων είναι αρκετά και σε διαφορετικά επίπεδα. Για να τα αποτυπώσουμε, δίνοντας μία πλήρη και εύληπτη εικόνα, μπορούμε να σχεδιάσουμε δύο θηκογράμματα. Το θηκόγραμμα είναι μια γραφική παράσταση που μας βοηθάει να απεικονίσουμε την κατανομή μιας ποσοτικής μεταβλητής στις κατηγορίες μιας ποιοτικής μεταβλητής. Από τη διάμεσο μπορούμε να πάρουμε μια ιδέα για την κεντρική τάση των δεδομένων, ενώ από το μήκος του (ολόκληρου, αλλά και του ορθογωνίου) μπορούμε να δούμε πόσο ποικίλλουν οι τιμές (μεταβλητότητα). Έτσι, λοιπόν, στο σχήμα 1 φαίνονται σχηματικά όσα περιγράψαμε προηγουμένως για τα δρομολόγια Μαΐου και Ιουνίου. Η μεταβλητότητα της διάρκειας του δρομολογίου τον Ιούλιο είναι μεγαλύτερη από τον μήνα Μάιο και για τα δρομολόγια «μέσης διάρκειας» (δηλαδή εντός ενδοτεταρτημοριακού εύρους), όπως φαίνεται από το μήκος του ορθογωνίου, αλλά και για όλα τα «τυπικά» δρομολόγια, όπως φαίνεται από το συνολικό μήκος των οριζοντίων γραμμών των θηκογραμμάτων. Επίσης, μπορούμε να εντοπίσουμε τις ακραίες τιμές της διάρκειας των δρομολογίων. Τον Ιούλιο έχουμε μόνο μία, ενώ τον Μάιο έχουμε τρεις ακραίες τιμές (από τις οποίες οι δύο είναι μεγαλύτερες από αυτή του Μαΐου). Ενδεχομένως να συνέβη κάτι έκτακτο σε εκείνα τα δρομολόγια. Λόγω των ακραίων τιμών, το εύρος της διάρκειας των δρομολογίων τον Μάιο είναι μεγαλύτερο, κάτι που φαίνεται και στα θηκογράμματα.

Κοιτώντας μόνο τα θηκογράμματα, φαίνεται τα «τυπικά» δρομολόγια τον Ιούλιο να κρατάνε λίγο περισσότερο και η διάρκειά τους να έχει λίγο μεγαλύτερη διασπορά.



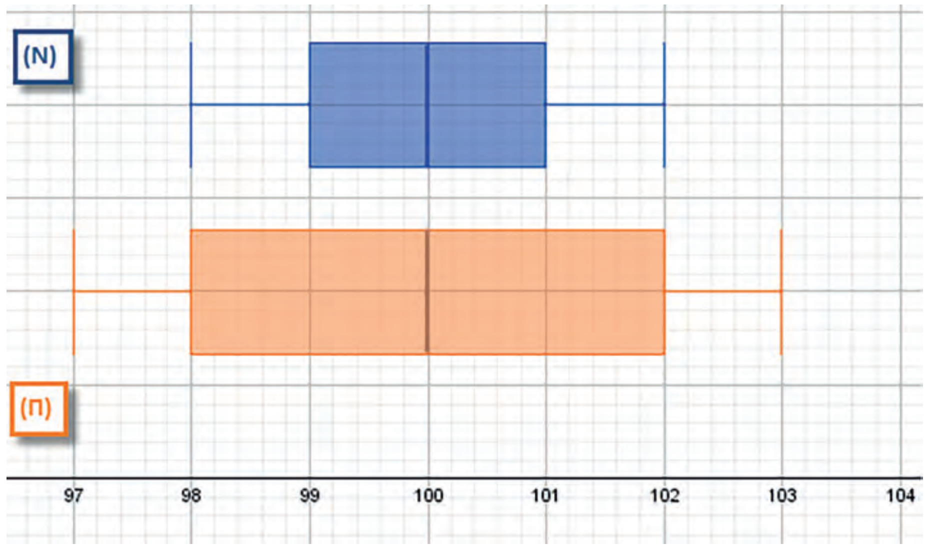

*Σχήμα 1: Θηκόγραμμα της διάρκειας ενός συγκεκριμένου δρομολογίου για τους μήνες Μάιο και Ιούλιο*

Αυτό, ενδεχομένως συμβαίνει γιατί τον Ιούλιο το λεωφορείο είναι περισσότερο γεμάτο και χρειάζεται περισσότερος χρόνος για την επιβίβαση/αποβίβαση, ενώ και η κίνηση στο νησί μπορεί να είναι μεγαλύτερη και να κάνει τη διάρκεια των δρομολογίων απρόβλεπτη. Ωστόσο, κοιτώντας τους μέσους χρόνους των δρομολογίων τους δύο μήνες, βλέπουμε ότι δε διαφέρουν πολύ. Άρα, αν συγκρίναμε μόνο τους μέσους χρόνους, χωρίς τα θηκογράμματα, δε θα είχαμε αυτή την εικόνα. Ο λόγος που οι μέσοι χρόνοι δε διαφέρουν πολύ είναι διότι επηρεάζονται από τις ακραίες τιμές και τον Μάιο είχαμε περισσότερες, όπως είπαμε.

## Εφαρμογές

### Εφαρμογή 1

Σ' ένα εργοστάσιο κατασκευής τροφίμων πρόκειται να αγοραστεί μια καινούργια μηχανή συσκευασίας (Ν) για να αντικαταστήσει την παλαιού τύπου μηχανή (Π). Έγινε μια δοκιμή για να ελεγχθεί η δυνατότητα παραγωγής της νέας μηχανής. Επελέγησαν 1000 πακέτα τροφίμων από κάθε μηχανή και μετρήθηκαν τα αντίστοιχα βάρη. Για τις δύο μηχανές δίνονται οι δειγματικές μέσες τιμές $\bar{x}_{\Pi} = \bar{x}_{N} = 100$ και τα αντίστοιχα θηκογράμματα τους.




*Σχήμα 2: Θηκόγραμμα του βάρους 1000 πακέτων τροφίμων από τις μηχανές τύπου (Π) και (Ν).*

Θα συστήνατε στον ιδιοκτήτη του εργοστασίου να αντικαταστήσει τη μηχανή (Π) με τη μηχανή (Ν); Δικαιολογήστε την απάντησή σας.

**Λύση**

Οι δειγματικές μέσες τιμές για τις δύο μηχανές τύπου (Π) και (Ν) είναι ίσες, όπως και οι δειγματικές διάμεσοι, όπως προκύπτει από το παραπάνω θηκόγραμμα. Αν, λοιπόν, δεν εξετάζαμε κάποιο άλλο στοιχείο, τότε δε θα είχε νόημα να συστήσουμε στον ιδιοκτήτη του εργοστασίου να αλλάξει τη μηχανή. Όμως, βλέπουμε από το θηκόγραμμα ότι η μεταβλητότητα των δεδομένων του δείγματος που πήραμε από τη μηχανή (Ν) είναι μικρότερη από τη μεταβλητότητα των δεδομένων του δείγματος που πήραμε από τη μηχανή (Π). Αυτό σημαίνει ότι για τα δύο δείγματα η μηχανή (Π) μας δίνει πακέτα τροφίμων που απέχουν περισσότερο από το μέσο βάρος των πακέτων της, σε σύγκριση με τη μηχανή (Ν). Συνεπώς, με δεδομένο ότι οι δειγματικές μέσες τιμές των πακέτων είναι ίσες για τις δύο μηχανές, θα συστήναμε στον ιδιοκτήτη του εργοστασίου να αλλάξει τη μηχανή, αν έχει σημασία η ακρίβεια του βάρους κάθε πακέτου.

## Εφαρμογή 2

Δύο μεγάλες εταιρείες, μία αμερικάνικη και μία ευρωπαϊκή, λαμβάνουν υπηρεσίες και από εξωτερικούς συνεργάτες εκτός από τους υπαλλήλους τους (π.χ. εκπόνηση μελετών, υπηρεσίες φύλαξης, ταχυδρομικές, κτλ). Στο πλαίσιο ενός μεγάλου διαγωνισμού είχαν πολλές τέτοιες συνεργασίες. Επιλέγουμε τυχαία δύο δείγματα αμοιβών εξωτερικών συνεργατών (ένα για κάθε εταιρεία), που δόθηκαν γι' αυτό το σκοπό. Παρακάτω δίνονται οι δειγματικές μέσες τιμές και οι δειγματικές τυπικές αποκλίσεις.

- **Α (αμερικάνικη εταιρεία):** $\bar{x}_A = 700$ δολάρια και $s_A = 80$ δολάρια
- **Ε (ευρωπαϊκή εταιρεία):** $\bar{x}_E = 550$ ευρώ και $s_E = 125$ ευρώ.

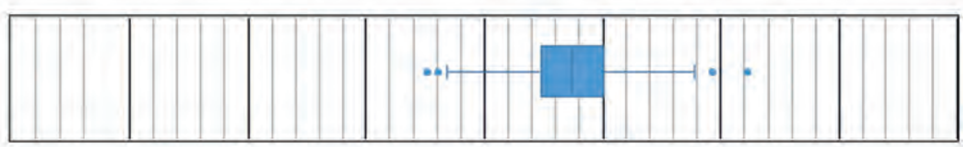
**α)** Να συγκρίνετε τους συντελεστές μεταβλητότητας των δύο δειγμάτων αμοιβών.

**β)** Οι δύο εταιρείες κατάφεραν να κερδίσουν στον διαγωνισμό και να αναλάβουν ένα έργο που θα τους αποφέρει αρκετά κέρδη. Για να επιβραβεύσουν τους συνεργάτες τους, η αμερικάνικη εταιρεία αποφάσισε να προσθέσει στην αμοιβή κάθε συνεργάτη 250 δολάρια και η ευρωπαϊκή εταιρεία να αυξήσει τη μηνιαία αμοιβή κάθε συνεργάτη κατά 40%. Να βρείτε τη νέα δειγματική μέση τιμή και τη νέα δειγματική τυπική απόκλιση των αμοιβών και για τις δύο εταιρείες.

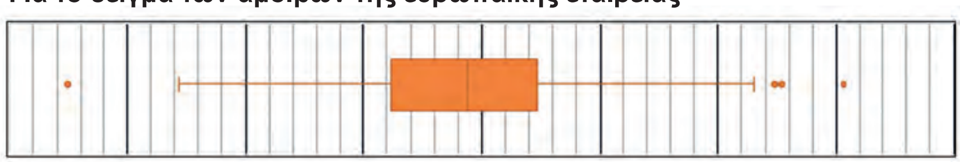
**γ)** Να συγκρίνετε τους συντελεστές μεταβλητότητας των δύο δειγμάτων αμοιβών μετά τις αυξήσεις.

**δ)** Παρακάτω φαίνονται τα θηκογράμματα των αμοιβών, μετά τις αυξήσεις, για τα δύο δείγματα της αμερικάνικης και της ευρωπαϊκής εταιρείας (σε δολάρια και ευρώ).

**Για το δείγμα των αμοιβών της αμερικάνικης εταιρείας**




**Για το δείγμα των αμοιβών της ευρωπαϊκής εταιρείας**




Για τα θηκογράμματα δίνονται, βοηθητικά, οι πίνακες:

| | Αμερικάνικη εταιρεία | Ευρωπαϊκή εταιρεία |
|---|---|---|
| $Q_1 \rightarrow$ | 894,8 | 647,9 |
| $\delta \rightarrow$ | 947,1 | 776,1 |
| $Q_3 \rightarrow$ | 1000,5 | 893,8 |
| Ακραίες τιμές | 720,7 , 704,3 , 1188,5 , 1243,6 | 97,7 , 1296,5 , 1307,35 , 1411,9 |

Πώς μπορείτε να χρησιμοποιήσετε τα θηκογράμματα για να συγκρίνετε τη μεταβλητότητα στις αμοιβές των δύο εταιρειών;

## Λύση

**α)** Οι συντελεστές μεταβλητότητας των αμοιβών των δύο εταιρειών, για τα δείγματα, είναι:

$$CV_A = \frac{s_A}{\bar{x}_A} = \frac{80}{700} = 11\% \text{ και } CV_E = \frac{s_E}{\bar{x}_E} = \frac{125}{550} = 23\%.$$

Ισχύει ότι $CV_A < CV_E$, οπότε το δείγμα από την αμερικάνικη εταιρεία είναι πιο ομοιογενές από το αντίστοιχο ευρωπαϊκό δείγμα.

**β)** Οι αμοιβές των συνεργατών της αμερικάνικης εταιρείας αυξάνονται κατά 250 δολάρια, οπότε η νέα μέση τιμή θα είναι $\bar{y}_A = \bar{x}_A + 250$, δηλαδή $\bar{y}_A = 950$ δολάρια. Η νέα τυπική απόκλιση θα είν αι $s'_A = s_A = 80$.

Οι αμοιβές των συνεργατών της ευρωπαϊκής εταιρείας θα αυξηθούν κατά 40%, οπότε η νέα μέση τιμή θα είναι $\bar{y}_E = 1,4 \cdot \bar{x}_E$, δηλαδή $\bar{y}_E = 770$ευρώ. Η νέα τυπική απόκλιση θα είναι $s'_E = 1,4 \cdot s_E$, δηλαδή $s'_E = 175$ ευρώ.

**γ)** Οι νέοι συντελεστές μεταβλητότητας είναι:
$CV'_A = \frac{s'_A}{\bar{y}_A} = \frac{80}{950} = 8\%$ και $CV'_E = \frac{s'_E}{\bar{y}_E} = \frac{1,4 \cdot s_E}{1,4 \cdot \bar{x}_E} = CV_E = 23\%$. Βλέπουμε ότι ισχύει η σχέση $CV'_A < CV'_E$ , καθώς ο συντελεστής μεταβλητότητας είναι ανεξάρτητος της νομισματικής μονάδας (δολάρια ή ευρώ).

**δ)** Κάτι που περιμένουμε να δούμε στα δύο θηκογράμματα είναι η μικρότερη μεταβλητότητα των αμοιβών του πρώτου δείγματος, δηλαδή της αμερικάνικης εταιρείας. Ωστόσο, για να συγκρίνουμε τα θηκογράμματα θα πρέπει οι αμοιβές να είναι υπολογισμένες στην ίδια νομισματική μονάδα. Θα προσαρμόσουμε το θηκόγραμμα της αμερικάνικης εταιρείας, χρησιμοποιώντας τη νομισματική ισοτιμία 1 δολάριο = 0,9 ευρώ για να το μετατρέψουμε.

Χρησιμοποιούμε τον βοηθητικό πίνακα της αμερικάνικης εταιρείας, πολλαπλασιάζοντας τους αριθμούς με 0,9. Παίρνουμε τα εξής αποτελέσματα:

| | Αμερικάνικη εταιρεία |
|---|---|
| $Q_1 \rightarrow$ | 805,32 |
| $\delta \rightarrow$ | 852,39 |
| $Q_3 \rightarrow$ | 900,45 |
| Ακραίες τιμές | 648,612 , 633,915 , 1119,114 , 1069,668 |

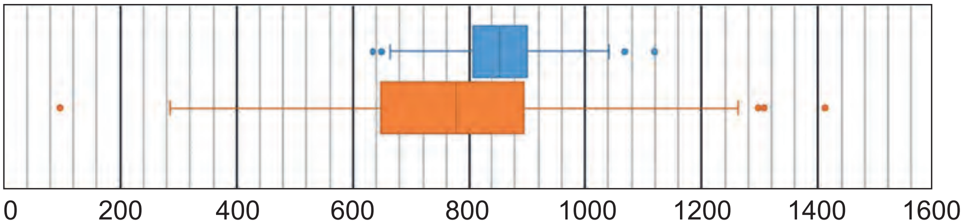
Υπολογίζουμε τα παρακάτω που μας χρειάζονται για να σχεδιάσουμε το νέο θηκόγραμμα:

$$Q = Q_3 - Q_1 = 900,45 - 805,32 = 95,13$$

$$Q_1 - 1,5Q = 805,32 - 1,5 \cdot 95,13 = 662,625$$

$$Q_3 + 1,5Q = 900,45 + 1,5 \cdot 95,13 = 1043,145$$

Το νέο θηκόγραμμα για το δείγμα αμοιβών της αμερικάνικης εταιρείας φαίνεται στο παρακάτω σχήμα, μαζί με το θηκόγραμμα της ευρωπαϊκής εταιρείας.



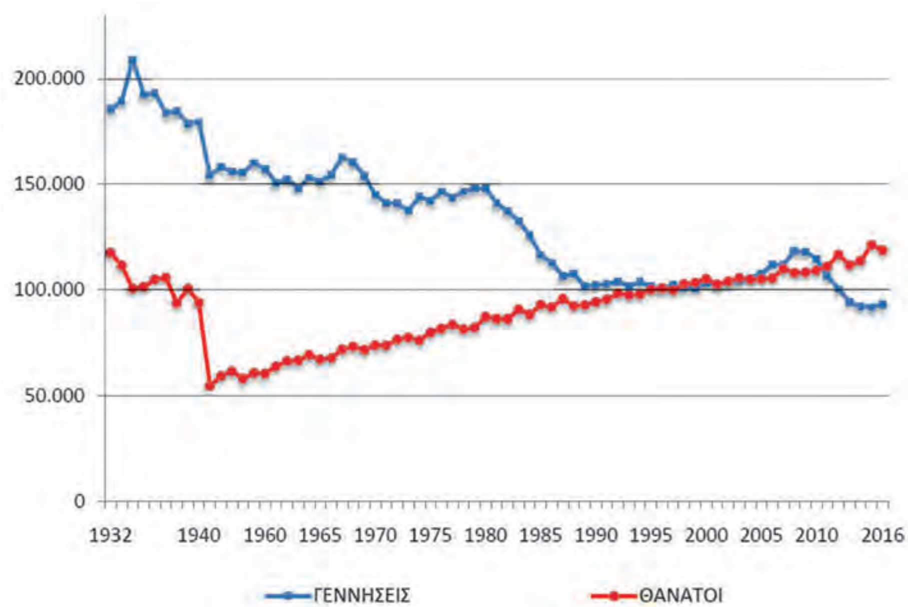

## Εφαρμογή 3

Στο χρονόγραμμα που δίνεται παρακάτω φαίνονται οι γεννήσεις ζώντων και οι θάνατοι στη χώρα μας από το 1932 έως και το 2016 (Πηγή: ΕΛ.ΣΤΑΤ: Διεύθυνση Στατιστικών Πληθυσμού & Αγοράς Εργασίας, Τμήμα Φυσικής & Μεταναστευτικής Κίνησης Πληθυσμού).

**α)** Ποια χρονιά είχαμε τη μεγαλύτερη αύξηση του πληθυσμού;

**β)** Ποια χρονιά ο αριθμός των θανάτων ξεπερνά για πρώτη φορά τον αριθμό των γεννήσεων στη χώρα μας;

**γ)** Τι παρατηρείτε την περίοδο 2005 – 2010 και τι συμβαίνει μετά το 2010;




### Λύση

**α)** Από το κοινό χρονόγραμμα δύο μεταβλητών (γεννήσεων και θανάτων) από το 1932 έως και το 2016, βλέπουμε ότι η μεγαλύτερη αύξηση πληθυσμού παρατηρήθηκε το 1934, δεδομένου ότι η διαφορά «γεννήσεις - θάνατοι» (στο σχήμα, η απόσταση του αριθμού των θανάτων από τον αριθμό των γεννήσεων το 1934) ήταν μεγαλύτερη από κάθε άλλη χρονιά. Από το αρχείο της ΕΛ.ΣΤΑΤ. μπορούμε να βρούμε και τους ακριβείς αριθμούς. Για το 1934 οι γεννήσεις ήταν 208.929 και οι θάνατοι 100.651 (διαφορά: 108.278) και ακολουθούν οι χρονιές 1955 (διαφορά: 99482) και 1959 (διαφορά: 99.347).

**β)** Ο αριθμός των θανάτων ξεπερνά για πρώτη φορά τον αριθμό των γεννήσεων στη χώρα μας το 1996. Ωστόσο, εκείνη τη χρονιά η διαφορά ήταν πολύ μικρή (διαφορά: -22). Αν εξαιρέσουμε το 1997, όπου είχαμε αύξηση των γεννήσεων, από το 1998 μέχρι και το 2003 οι θάνατοι ήταν περισσότεροι.

**γ)** Την περίοδο 2005 – 2010 καταγράφηκε αύξηση του πληθυσμού, όπως φαίνεται και από το χρονόγραμμα, ενώ από το 2010 και μέχρι το 2016 είχαμε σημαντική μείωση του πληθυσμού. Αξίζει να σημειωθεί ότι οι διαφορές από το 2011 έως και το 2016 ήταν: -4.671, -16.297, -17.660, -21.591, -29.336 και -25.887, αντίστοιχα. Το γεγονός αυτό θα μπορούσε να αποδοθεί κατά κύριο λόγο στην οικονομική κρίση που ξεκίνησε στην Ελλάδα το 2010.

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Στους επόμενους πίνακες δίνονται οι βαθμολογίες των μαθητών/τριών δύο τμημάτων της Β΄ τάξης ενός γενικού λυκείου σε μια γραπτή αξιολόγηση της Άλγεβρας:

| Τμήμα $B_1$ | | | |
|---|---|---|---|
| 20 | 17 | 14 | 10 |
| 20 | 17 | 13 | 9 |
| 19 | 16 | 12 | 9 |
| 19 | 16 | 11 | 8 |
| 17 | 15 | 10 | 8 |

| Τμήμα $B_2$ | | | |
|---|---|---|---|
| 20 | 19 | 14 | 11 |
| 20 | 19 | 14 | 10 |
| 20 | 18 | 14 | 9 |
| 20 | 15 | 13 | 9 |
| 19 | 15 | 12 | 8 |

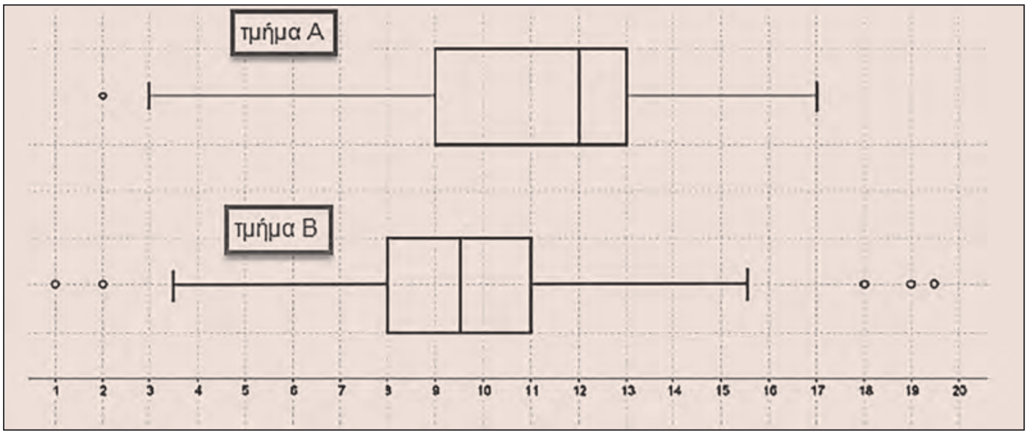
**α)** Να υπολογίσετε τη μέση τιμή και την τυπική απόκλιση της βαθμολογίας των μαθητών/τριών σε κάθε τμήμα. Ποια πρώτη εικόνα σας δίνουν τα αποτελέσματα των παραπάνω στατιστικών μέτρων για την επίδοση κάθε τμήματος;

**β)** Να βρείτε τους συντελεστές μεταβλητότητας (CV) και να συγκρίνετε τα δύο τμήματα ως προς την ομοιογένειά τους.

**γ)** Ο καθηγητής θέλει να δώσει βραβείο στους μαθητές/μαθήτριες κάθε τμήματος που πήραν βαθμό μεγαλύτερο ή ίσο από το 75% των μαθητών/τριών του τμήματος και να δώσει επιπλέον εργασία για το σπίτι στους μαθητές/μαθήτριες που πήραν βαθμό μικρότερο ή ίσο από το 25% των μαθητών/τριών του τμήματος. Να βρείτε τους βαθμούς των μαθητών/τριών που θα βραβευτούν και τους βαθμούς των μαθητών/τριών που θα πάρουν επιπλέον εργασία.

**δ)** Να κατασκευάσετε τα θηκογράμματα για κάθε τμήμα.

**2)** Τα παρακάτω θηκογράμματα παρουσιάζουν τους βαθμούς των μαθητών/τριών δύο τμημάτων Α και Β σε έναν μαθηματικό διαγωνισμό.




**α)** Να βρείτε ποιο από τα δύο τμήματα έχει το μεγαλύτερο εύρος βαθμών.

**β)** Να βρείτε ποιο από τα δύο τμήματα έχει το μεγαλύτερο ενδοτεταρτημοριακό εύρος βαθμών;

**γ)** Σε ποιο από τα δύο τμήματα φαίνεται να είναι πιο συμμετρική γύρω από τη διάμεσο η κατανομή των βαθμών;

**δ)** Να γράψετε μια μικρή αναφορά για το ποιο τμήμα θα μπορούσε να είναι το καλύτερο.

**ε)** Να βρείτε το τμήμα και τη βαθμολογία των δύο μαθητών/τριών με τον καλύτερο βαθμό.

**3)** Ο επόμενος πίνακας παρουσιάζει τον αριθμό των απουσιών που έκαναν οι μαθητές ενός τμήματος τους μήνες Οκτώβριο και Νοέμβριο:

| | **Απουσίες μαθητών ενός τμήματος** | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Οκτώβριος | 14 | 7 | 0 | 19 | 21 | 7 | 3 | 0 | 0 | 25 | 2 | 9 | 8 | 7 | 14 | 20 | 0 |
| Νοέμβριος | 7 | 8 | 2 | 2 | 1 | 23 | 14 | 7 | 0 | 14 | 4 | 7 | 5 | 0 | 1 | 8 | 0 |

Να κατασκευάσετε το θηκόγραμμα των απουσιών των μαθητών του τμήματος για κάθε μήνα χωριστά και να συγκρίνετε τις απουσίες των μαθητών.

**4)** Για δύο τύπους μπαταριών Α και Β επιλέχθηκαν δύο δείγματα μεγέθους 5 το καθένα. Οι χρόνοι ζωής των μπαταριών για το κάθε δείγμα σε χιλιάδες ώρες δίνονται στον επόμενο πίνακα:

| Μπαταρία τύπου Α | Μπαταρία τύπου Β |
|---|---|
| 20 | 26 |
| 26 | 32 |
| 24 | 19 |
| 22 | 20 |
| 18 | 23 |

**α)** Να βρείτε τη δειγματική μέση τιμή της διάρκειας ζωής μιας μπαταρίας τύπου Α και μιας μπαταρίας τύπου Β.

**β)** Με βάση το παραπάνω δείγμα και το γεγονός ότι μια μπαταρία τύπου Α στοιχίζει 38 ευρώ, ποιου τύπου μπαταρία θα προτιμήσετε αν μια μπαταρία τύπου Β στοιχίζει:

**i)** 40 ευρώ **ii)** 42 ευρώ

Να δικαιολογήσετε την απάντησή σας σε κάθε μία από τις περιπτώσεις **i)** και **ii)**.

**γ)** Να βρείτε τις δειγματικές τυπικές αποκλίσεις $s_A$ και $s_B$ της διάρκειας ζωής των δύο τύπων μπαταριών.

**δ)** Να βρείτε ποιο από τα δύο παραπάνω δείγματα μπαταριών τύπου Α και Β παρουσιάζει τη μεγαλύτερη ομοιογένεια ως προς τη διάρκεια ζωής.

**5)** Στις 12μ.μ. η θερμοκρασία (σε βαθμούς Κελσίου) στη Λαμία και στη Θεσσαλονίκη το τελευταίο δεκαήμερο του Μαρτίου ήταν:

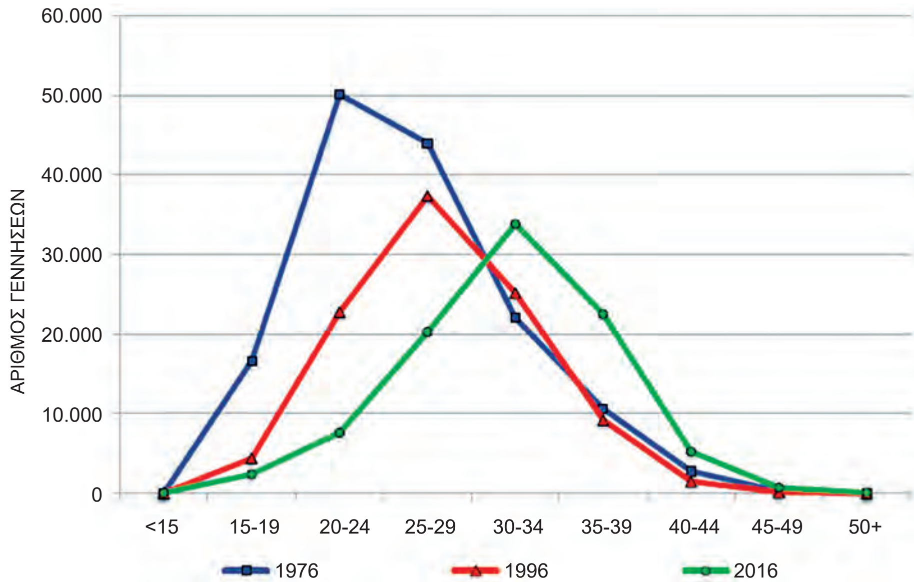
| | **Θερμοκρασία σε βαθμούς Κελσίου** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Λαμία (Λ)** | **20** | **18** | **20** | **17** | **18** | **17** | **16** | **17** | **16** | **10** |
| **Θεσσαλονίκη (Θ)** | **18** | **16** | **17** | **15** | **16** | **12** | **16** | **17** | **20** | **22** |

**α)** Να βρείτε τη μέση, τη διάμεσο και την επικρατούσα θερμοκρασία των δειγμάτων της Λαμίας και της Θεσσαλονίκης.

**β)** Αν η δειγματική τυπική απόκλιση (σε βαθμούς Κελσίου) για τη Λαμία και τη Θεσσαλονίκη είναι $s_Λ = 2{,}66$ και $s_Θ = 2{,}59$ αντίστοιχα, να δικαιολογήσετε σε ποια από τα δύο δείγματα οι τιμές της θερμοκρασίας έχουν μεγαλύτερη διασπορά.

**γ)** Εκ των υστέρων διαπιστώθηκε ότι το θερμόμετρο που χρησιμοποιήθηκε για τη μέτρηση της θερμοκρασίας στη Λαμία παρουσίαζε, λόγω κατασκευαστικού λάθους, αυξημένη θερμοκρασία κατά 5 βαθμούς. Αφού υπολογίσετε τις σωστές θερμοκρασίες της Λαμίας, να βρείτε σε ποια πόλη από τις δύο το συγκεκριμένο δεκαήμερο οι τιμές της θερμοκρασίας έχουν μεγαλύτερη ομοιογένεια. Να δικαιολογήσετε την απάντησή σας.

**6)** Στο επόμενο διάγραμμα φαίνονται τα πολύγωνα συχνοτήτων των γεννήσεων ζώντων κατά ομάδες ηλικιών της μητέρας για τα έτη 1976, 1996 και 2016 (Πηγή: ΕΛ.ΣΤΑΤ: Διεύθυνση Στατιστικών Πληθυσμού & Αγοράς Εργασίας, Τμήμα Φυσικής & Μεταναστευτικής Κίνησης Πληθυσμού).




**α)** Σε ποια ηλικία των μητέρων έχουμε τις περισσότερες γεννήσεις το 1976, το 1996 και το 2016; Μπορείτε να το δικαιολογήσετε;

**β)** Να σχολιάσετε τον αριθμό γεννήσεων κατά ηλικιακή ομάδα των μητέρων για το 2016.

**γ)** Σε ποιες ηλικιακές ομάδες των μητέρων φαίνεται ο αριθμός των γεννήσεων το 2016 να ξεπερνάει τους αντίστοιχους αριθμούς για τα έτη 1996 και 1976; Γιατί πιστεύετε ότι συνέβη αυτό;

# Πρόσθετο Υλικό

## Θέματα για διερεύνηση

**1)** Ένα από τα σοβαρά προβλήματα, που απασχολεί ολοένα και περισσότερους στην ελληνική κοινωνία, είναι η δημογραφική συρρίκνωση του συνολικού πληθυσμού της Ελλάδας. Σε συνδυασμό με την άσκηση 6 και τον επόμενο πίνακα που δίνεται, να γράψετε μια μικρή αναφορά την οποία πρόκειται να παρουσιάσετε στο σχολείο σας (ή με τα μέσα κοινωνικής δικτύωσης), εξηγώντας την πορεία των γεννήσεων και των θανάτων από το 1932 έως το 2016. Πώς μπορείτε να πείσετε την πολιτεία να λάβει κάποια μέτρα; Αναζητήστε τα στοιχεία για τη φετινή χρονιά στην Ελληνική Στατιστική Αρχή (Δημογραφικά Στοιχεία – Γεννήσεις) για να ενισχύσετε την επιχειρηματολογία σας. Η εργασία αυτή μπορεί να υλοποιηθεί στα πλαίσια κάποιας διαθεματικής – συνθετικής εργασίας. Τα πλήρη δεδομένα των ερευνών σε μορφή .xls θα τα βρείτε στον παραπάνω σύνδεσμο.

| Έτος | Γεννήσεις | Θάνατοι | Έτος | Γεννήσεις | Θάνατοι |
|---|---|---|---|---|---|
| **1932** | **185.523** | **117.593** | **2000** | **103.274** | **105.170** |
| **1940** | **179.500** | **93.830** | **2005** | **107.545** | **105.091** |
| **1960** | **157.239** | **60.563** | **2010** | **114.766** | **109.084** |
| **1965** | **151.448** | **67.269** | **2011** | **106.428** | **111.099** |
| **1970** | **144.928** | **74.009** | **2012** | **100.371** | **116.668** |
| **1975** | **142.273** | **80.077** | **2013** | **94.134** | **111.794** |
| **1980** | **148.134** | **87.282** | **2014** | **92.149** | **113.740** |
| **1985** | **116.481** | **92.886** | **2015** | **91.847** | **121.183** |
| **1990** | **102.229** | **94.152** | **2016** | **92.898** | **118.785** |
| **1995** | **101.495** | **100.158** | | | |

*Πίνακας 2: Γεννήσεις και Θάνατοι από το 1932 – 2016*

**2)** Στον επόμενο πίνακα παρουσιάζονται οι αριθμοί θρησκευτικών και πολιτικών γάμων που έγιναν στην Ελλάδα από το 1991 έως το 2016. Επιπλέον, προστέθηκαν από το 2009 και μετά τα σύμφωνα συμβίωσης (Πηγή: ΕΛ.ΣΤΑΤ: Γάμοι). Να γράψετε μια μικρή αναφορά σχετικά με την εξέλιξη του θρησκευτικού και πολιτικού γάμου στη χώρα μας. Έπαιξαν κάποιο ρόλο μετά το 2009 τα σύμφωνα συμβίωσης; Ενισχύστε την επιχειρηματολογία σας αντλώντας πληροφορίες από τον πίνακα, δηλαδή:

- Βρείτε τα ποσοστά θρησκευτικών, πολιτικών γάμων και σύμφωνων συμβίωσης, ανά έτος. Συγκρίνετε τα ποσοστά κυρίως μετά το 2009. Τι παρατηρείτε τις τελευταίες χρονιές 2015 και 2016;

- Βρείτε τις μεταβολές των θρησκευτικών και πολιτικών γάμων και σύμφωνων συμβίωσης ανάμεσα στα έτη 2013-2014 και στα έτη 2015-2016.
- Κατασκευάστε τα αντίστοιχα χρονογράμματα.

| ΕΤΟΣ | ΑΠΟΛΥΤΟΙ ΑΡΙΘΜΟΙ | | | |
|---|---|---|---|---|
| | Σύνολο | Θρησκευτικοί | Πολιτικοί | Σύμφωνα Συμβίωσης |
| 1991 | 65.568 | 59.710 | 5.858 | 0 |
| 2001 | 58.491 | 48.087 | 10.404 | 0 |
| 2009 | 59.212 | 34.375 | 24.837 | 161 |
| 2010 | 56.338 | 30.327 | 26.011 | 180 |
| 2011 | 55.099 | 28.472 | 26.627 | 185 |
| 2012 | 49.705 | 23.980 | 25.725 | 314 |
| 2013 | 51.256 | 25.624 | 25.632 | 581 |
| 2014 | 53.105 | 26.152 | 26.953 | 1573 |
| 2015 | 53.672 | 26.419 | 27.253 | 2611 |
| 2016 | 49.632 | 23.778 | 25.854 | 3799 |

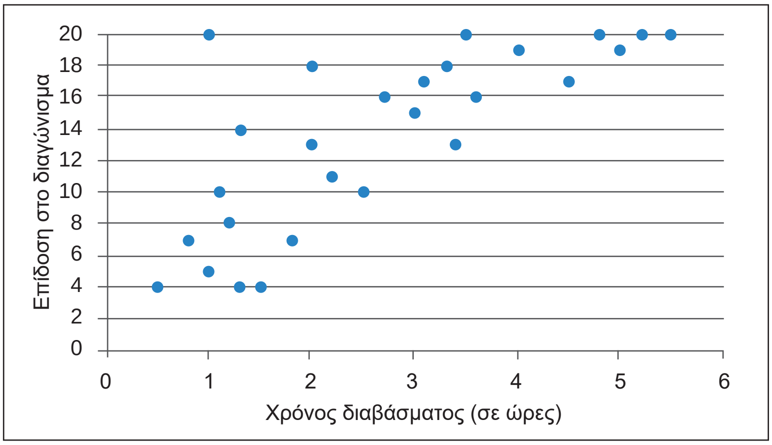
*Πίνακας 3: Αριθμοί θρησκευτικών, πολιτικών γάμων και σύμφωνων συμβίωσης από το 1991 – 2016*

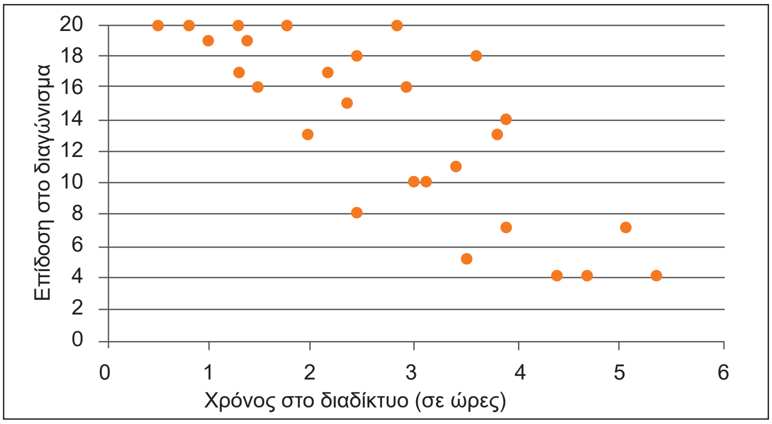
**3)** Στα πλαίσια μιας συνθετικής εργασίας συλλέξτε δεδομένα που αφορούν πληροφορίες για τους συμμαθητές σας. Συντάξτε ένα ερωτηματολόγιο με ερωτήσεις που αφορούν τα βασικά στοιχεία των συμμαθητών σας (φύλο, τάξη, ομάδα προσανατολισμού) και ερωτήσεις (ύψος, βάρος, ώρες διαβάσματος, ώρες στο διαδίκτυο, βαθμοί σε διάφορα μαθήματα). Φροντίστε οι ερωτήσεις να είναι σαφώς διατυπωμένες και οι πιθανές απαντήσεις να μην επικαλύπτονται. Χρησιμοποιώντας αυτά που μάθατε στη συγκεκριμένη παράγραφο, γράψτε μια σύντομη ερευνητική έκθεση παρουσιάζοντας τα αποτελέσματά σας.

# ΕΝΟΤΗΤΑ 2.7 ΓΡΑΜΜΙΚΗ ΣΥΣΧΕΤΙΣΗ ΠΟΣΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΚΑΙ ΔΙΑΓΡΑΜΜΑΤΑ ΔΙΑΣΠΟΡΑΣ

## Διερεύνηση

Τα επόμενα διαγράμματα παρουσιάζουν την επίδοση 27 μαθητών ενός τμήματος της Β΄ τάξης του Γενικού Λυκείου στην Άλγεβρα σε σχέση με τον χρόνο που αφιέρωσαν τις προηγούμενες δύο μέρες για διάβασμα και ενασχόληση στο διαδίκτυο.




**α)** Ποιες είναι οι μεταβλητές του πρώτου και δεύτερου διαγράμματος; Ποια από τις δύο μπορεί να θεωρηθεί ως ανεξάρτητη και ποια ως εξαρτημένη από την άλλη;

**β)** Να συγκρίνετε τη σχέση των μεταβλητών «επίδοση στο διαγώνισμα» και «χρόνος διαβάσματος» με τη σχέση των μεταβλητών «επίδοση στο διαγώνισμα» και «χρόνος στο διαδίκτυο».

**γ)** Να χαράξετε «με το μάτι» μια ευθεία που να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Για ποιον λόγο θα χρησιμοποιούσατε μια ευθεία του **γ)** ερωτήματος;

## Βασικές Μαθηματικές Έννοιες – Ιδέες – Διεργασίες

Έχουμε δει μέχρι τώρα ότι η σχέση δύο ποιοτικών μεταβλητών περιγράφεται με τη βοήθεια των πινάκων συνάφειας και των κατάλληλων ραβδογραμμάτων, ενώ υπολογίζοντας τα μέτρα θέσης και μεταβλητότητας μιας ποσοτικής μεταβλητής στις κατηγορίες μιας ποιοτικής μπορούμε να συγκρίνουμε ομάδες δεδομένων. Τι συμβαίνει εάν έχουμε δύο ποσοτικές μεταβλητές; Πώς μπορούμε να περιγράψουμε τη σχέση δύο ποσοτικών μεταβλητών; Για παράδειγμα:

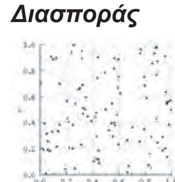
- Η ηλικία και το βάρος ενός παιδιού έχουν κάποια θετική συσχέτιση μεταξύ τους, με την έννοια ότι όσο πιο μεγάλο είναι το παιδί τόσο μεγαλύτερο βάρος θα έχει.
- Η διάρκεια ζωής των ζώντων οργανισμών σε μια περιοχή και ένας δείκτης μόλυνσης της περιοχής έχουν αρνητική συσχέτιση μεταξύ τους, με την έννοια ότι όσο πιο μεγάλος είναι ο δείκτης μόλυνσης της περιοχής τόσο μικρότερη είναι η διάρκεια ζωής των οργανισμών που ζουν στην περιοχή.
- Το ύψος των αποδοχών των υπαλλήλων μιας εταιρείας δε συσχετίζεται με το βάρος τους.

Στην ενότητα αυτή θα εξετάσουμε μόνο την περίπτωση της γραμμικής συσχέτισης μεταξύ δύο ποσοτικών μεταβλητών. Για παράδειγμα, καταγράφουμε το ύψος (X σε cm) και το βάρος (Y σε kg) 18 μαθητών της Β΄ Λυκείου. Αυτό που μας ενδιαφέρει είναι τι συμβαίνει με το βάρος των παιδιών όταν αλλάζει το ύψος τους ή τι συμβαίνει με το ύψος των παιδιών όταν αλλάζει το βάρος τους. Αν μας ενδιαφέρει η πρώτη περίπτωση, θεωρούμε ως ανεξάρτητη μεταβλητή το ύψος των μαθητών και ως εξαρτημένη (από το ύψος) μεταβλητή το βάρος των μαθητών. Στη δεύτερη περίπτωση συμβαίνει το αντίθετο.

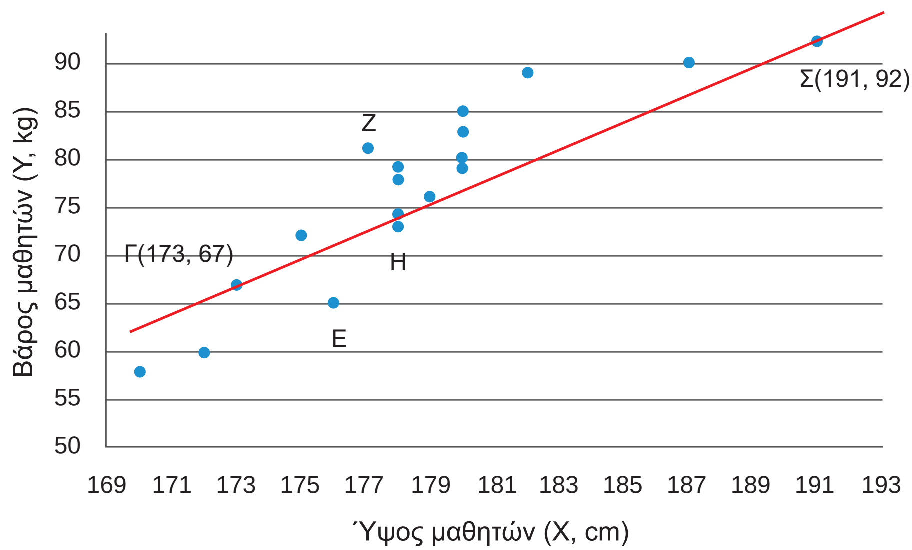
| Μαθητής | Ύψος (X) | Βάρος (Y) | Μαθητής | Ύψος (X) | Βάρος (Y) |
|---|---|---|---|---|---|
| Α | 170 | 58 | Κ | 178 | 79 |
| Β | 172 | 60 | Λ | 179 | 76 |
| Γ | 173 | 67 | Μ | 180 | 79 |
| Δ | 175 | 72 | Ν | 180 | 80 |
| Ε | 176 | 65 | Ξ | 180 | 83 |
| Ζ | 177 | 81 | Ο | 180 | 85 |
| Η | 178 | 73 | Π | 182 | 89 |
| Θ | 178 | 74 | Ρ | 187 | 90 |
| Ι | 178 | 78 | Σ | 191 | 92 |

*Πίνακας 1: Ύψη (X) και Βάρη (Y) 18 μαθητών της Β΄ Λυκείου*

***Διάγραμμα Διασποράς***



Για κάθε άτομο έχουμε δύο μετρήσεις: το ύψος και το βάρος του. Το δείγμα μας δηλαδή αποτελείται από τα ζεύγη τιμών $(x_i , y_i)$ , $i = 1,2,\dots,18$ των συνεχών μεταβλητών Χ και Υ. Η γραφική παράσταση των σημείων σε ένα ορθογώνιο σύστημα αξόνων ονομάζεται **διάγραμμα διασποράς** (scatterplot) των μεταβλητών Χ και Υ. Στο διάγραμμα παρατηρούμε μια διασπορά των σημείων που αντιστοιχούν στους μαθητές που εξετάζουμε. Από το διάγραμμα διασποράς μπορούμε να εντοπίσουμε την ύπαρξη συσχέτισης που ενδεχομένως να υπάρχει μεταξύ των μεταβλητών που εξετάζουμε. Η γενική εικόνα που παίρνουμε από το διάγραμμα διασποράς είναι ότι οι ψηλότεροι μαθητές είναι συνήθως και βαρύτεροι. Για παράδειγμα, ο μαθητής Ζ είναι ψηλότερος και βαρύτερος από τον μαθητή Ε, ο Η είναι ψηλότερος και βαρύτερος από τον Ε, αλλά υπάρχουν και εξαιρέσεις, όπως ο Η είναι ψηλότερος από τον Ζ αλλά ο Ζ είναι βαρύτερος από τον Η. Το άπλωμα των σημείων στο διάγραμμα θα μπορούσαμε να πούμε ότι γίνεται κοντά σε μια νοητή ευθεία γραμμή, οπότε η σχέση θεωρείται γραμμική.




*Σχήμα 1: Διάγραμμα Διασποράς Ύψους και Βάρους 18 μαθητών της Β΄ λυκείου και ευθεία προσαρμοσμένη «με το μάτι» για τα δεδομένα του πίνακα 1*

***Συντελεστής γραμμικής συσχέτισης των ποσοτικών μεταβλητών Χ και Υ***

Ένα μέτρο που μας δίνει το μέγεθος της γραμμικής σχέσης των δύο ποσοτικών μεταβλητών Χ και Υ είναι ο συντελεστής γραμμικής συσχέτισης των Χ και Υ (linear correlation coefficient ή Pearson's correlation coefficient). Ο συντελεστής γραμμικής συσχέτισης δύο ποσοτικών μεταβλητών Χ και Υ ορίζεται με βάση ένα δείγμα ν ζευγών παρατηρήσεων $(x_i , y_i)$ , i = 1,2,…,ν, συμβολίζεται με r (X , Y) ή απλά r και δίνεται από τον τύπο:

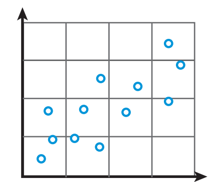
$$r = \frac{\sum_{i=1}^{\nu} x_i y_i - \nu \bar{x}\,\bar{y}}{\nu s_X s_Y}, \qquad (1)$$

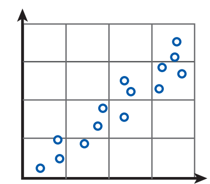
όπου $\bar{x}, s_X$ η δειγματική μέση τιμή και δειγματική τυπική απόκλιση της μεταβλητής Χ και $\bar{y}, s_Y$ η δειγματική μέση τιμή και δειγματική τυπική απόκλιση της μεταβλητής Υ. Ωστόσο, σε πραγματικές μελέτες με πολλά δεδομένα, η χρήση λογιστικού φύλλου ή κάποιου στατιστικού λογισμικού κρίνεται απαραίτητη.

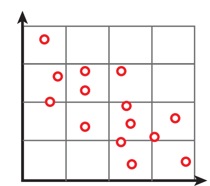
Ο συντελεστής γραμμικής συσχέτισης δύο ποσοτικών μεταβλητών Χ και Υ είναι καθαρός αριθμός, δηλαδή δεν εκφράζεται σε συγκεκριμένες μονάδες μέτρησης, επομένως είναι ανεξάρτητος των χρησιμοποιούμενων μονάδων μέτρησης των μεταβλητών Χ και Υ. Ισχύει πάντοτε:

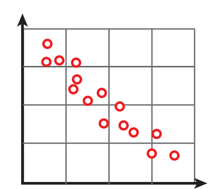
$$\boxed{-1 \leq r \leq 1}$$

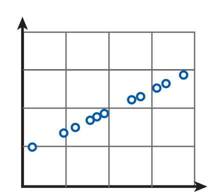
Πιο συγκεκριμένα όταν:

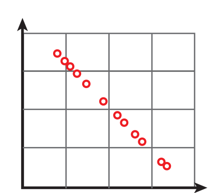

r = 0,50

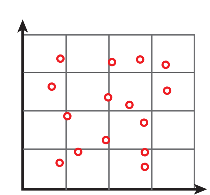

r = 0,90


r = -0,50


r = -0,90


r = 1


r = -1


r = 0

- $0 < r < 1$, τότε οι Χ και Υ είναι **θετικά γραμμικά συσχετισμένες**. Αυτό σημαίνει ότι, όταν οι τιμές της μεταβλητής Χ αυξάνονται, οι τιμές της Υ τείνουν να αυξάνονται. Όσο ο συντελεστής πλησιάζει το 1 τόσο πιο συγκεντρωμένες γύρω από μια νοητή ευθεία είναι οι παρατηρήσεις. Επομένως, τόσο πιο ισχυρή είναι η θετική γραμμική συσχέτιση.
- $-1 < r < 0$, τότε οι Χ και Υ είναι **αρνητικά γραμμικά συσχετισμένες**. Αυτό σημαίνει ότι, όταν οι τιμές της μεταβλητής Χ αυξάνονται, οι τιμές της Υ τείνουν να μειώνονται. Όσο ο συντελεστής πλησιάζει το -1 τόσο πιο συγκεντρωμένες γύρω από μια νοητή ευθεία είναι οι παρατηρήσεις. Επομένως, τόσο πιο ισχυρή είναι η αρνητική γραμμική συσχέτιση.
- $r = 1$, τότε οι Χ και Υ είναι **τέλεια θετικά γραμμικά συσχετισμένες** και όλα τα σημεία βρίσκονται πάνω σε μια ευθεία με θετική κλίση.
- $r = -1$, τότε οι Χ και Υ είναι **τέλεια αρνητικά γραμμικά συσχετισμένες** και όλα τα σημεία βρίσκονται πάνω σε μια ευθεία με αρνητική κλίση.
- $r = 0$, τότε οι Χ και Υ είναι **γραμμικά ασυσχέτιστες, χωρίς αυτό να σημαίνει ότι δεν μπορεί να έχουν κάποια άλλη σχέση, μη γραμμική**.

Από τα δεδομένα του παραδείγματος με τη χρήση αριθμομηχανής ή λογιστικού φύλλου έχουμε:

$$\sum_{i=1}^{18} x_i y_i = 247324, \bar{x} = 178,56, \bar{y} = 76,72, s_X = 4,809, s_Y = 9,473,$$

οπότε ο συντελεστής γραμμικής συσχέτισης του ύψους (Χ) και του βάρους (Υ) 18 μαθητών της Β Λυκείου είναι r = 0,90, που σημαίνει ότι το ύψος και το βάρος των μαθητών είναι ισχυρά θετικά γραμμικά συσχετισμένες μεταβλητές.

***Γραμμική συσχέτιση ποσοτικών μεταβλητών και αιτιότητα***

Στο σημείο αυτό θα υπενθυμίσουμε ότι η ισχυρή συσχέτιση μεταξύ δύο μεταβλητών δε συνεπάγεται υποχρεωτικά μια αιτιολογική σχέση ανάμεσα στις δύο μεταβλητές. Υπάρχει περίπτωση η αλλαγή της μεταβλητής X να προκαλεί άμεση αλλαγή της μεταβλητής Y. Αλλά πολύ συχνά οι αλλαγές των δύο μεταβλητών X και Y οφείλονται σε κάποιες άλλες μεταβλητές ή σε κάποιους αστάθμητους παράγοντες. Για παράδειγμα, το μέγεθος παπουτσιού που φορά ένας μαθητής και το μέγεθος των αρχείων που αποθηκεύονται στον υπολογιστή του, ενώ έχουν εξαιρετικά μεγάλη γραμμική συσχέτιση, προφανώς δεν έχουν αιτιώδη σχέση, αλλά επηρεάζονται από έναν τρίτο παράγοντα, την ηλικία (συγχυτικός παράγοντας).

***Χάραξη ευθείας στο διάγραμμα διασποράς «με το μάτι»***

Τα σημεία $(x_i, y_i)$, $i=1,2,\ldots,18$ είναι συγκεντρωμένα γύρω από μια ευθεία, όπως φαίνεται στο σχήμα 1. Στην πραγματικότητα, σχεδόν ποτέ δεν υπάρχει ευθεία που να περνάει απ' όλα τα σημεία των δεδομένων. Για παράδειγμα, στο πρόβλημά μας με το ύψος και το βάρος των μαθητών, η αλήθεια είναι ότι η μεταβλητή Y (το βάρος των μαθητών) μπορεί να επηρεάζεται και από άλλους παράγοντες εκτός της X (το ύψος των μαθητών), όπως, π.χ.. τις διατροφικές συνήθειες. Γι' αυτό μαθητές με το ίδιο ύψος μπορεί να έχουν διαφορετικό βάρος.

Ψάχνουμε, λοιπόν, για εκείνη την ευθεία που περνάει κοντά από τα περισσότερα σημεία των δεδομένων. Όπως γνωρίζουμε, η εξίσωση μιας ευθείας δίνεται από τη σχέση:

$$y = \alpha+\beta x, \qquad (2)$$

όπου α και β είναι παράμετροι τις οποίες θέλουμε να υπολογίσουμε ή, όπως λέμε, «να εκτιμήσουμε», έτσι ώστε η ευθεία που θα προκύψει από τη (2) να περιγράφει την αναμενόμενη τιμή y της μεταβλητής Y, όταν η τιμή της μεταβλητής X είναι ίση με x. Η παράμετρος α μας δίνει το σημείο (0, α) όπου η ευθεία αυτή τέμνει τον άξονα $y'y$ , ενώ η παράμετρος β παριστάνει τον συντελεστή διεύθυνσης της ευθείας. Ο πιο εύκολος τρόπος χάραξης της ευθείας είναι αυτός που γίνεται «με το μάτι». Μια τέτοια ευθεία έχουμε φέρει στο διάγραμμα διασποράς του σχήματος 1. Για να βρούμε τα α και β επιλέγουμε δύο σημεία, έστω τα Γ(173,67) και Σ(191,92) πάνω στην ευθεία που φέραμε «με το μάτι» και καταλήγουμε στην εξίσωση:

$$y = -173{,}28+1{,}39x \qquad (3)$$

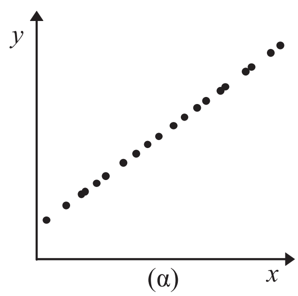
Επομένως, η ευθεία που κατά τη γνώμη μας προσαρμόζεται καλύτερα στα σημεία του διαγράμματος διασποράς διέρχεται από το σημείο (0,-173,28) και έχει συντελεστή διεύθυνσης 1,39.

Την εξίσωση της ευθείας (3) μπορούμε να τη χρησιμοποιήσουμε για να εκτιμήσουμε μια τιμή της μεταβλητής Y, από κάποια τιμή της μεταβλητής X που ανήκει στο εύρος τιμών της τελευταίας. Αυτό συμβαίνει, διότι –για τιμές μικρότερες ή μεγαλύτερες από αυτές που παρατηρήθηκαν– δεν είμαστε σίγουροι ότι η γραμμική σχέση παραμένει. Για παράδειγμα: ένας μαθητής με ύψος 1,81 εκτιμάται ότι θα έχει βάρος ίσο με $y = -173{,}28+1{,}39\cdot 181 = 78{,}1$ κιλά.

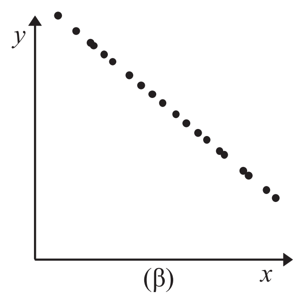
## Εφαρμογές

### Εφαρμογή 1

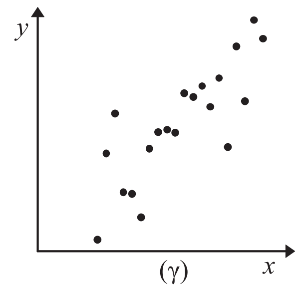
**A)** Για δύο ποσοτικές μεταβλητές X και Y να αντιστοιχίσετε τους συντελεστές γραμμικής συσχέτισης
$r_1 = -1$, $r_2 = -0{,}8$, $r_3 = -0{,}2$, $r_4 = 0$, $r_5 = 0{,}2$, $r_6 = 0{,}8$ και $r_7 = 1$ , με τα αντίστοιχα διαγράμματα διασποράς.



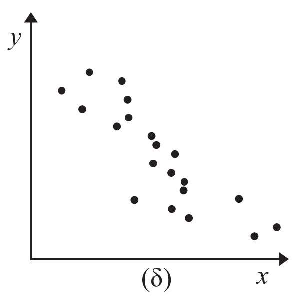
(α)



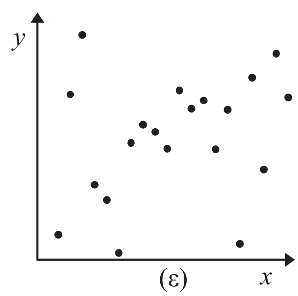
(β)



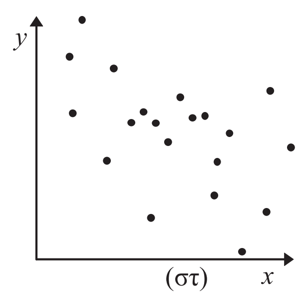
(γ)



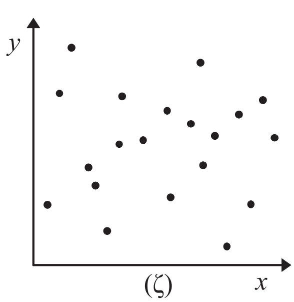
(δ)



(ε)

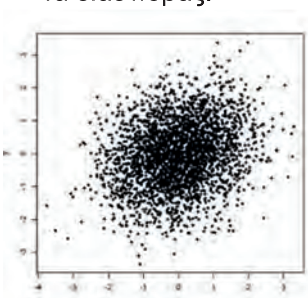


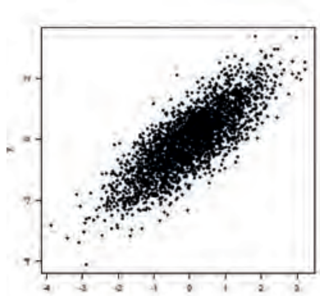
(στ)

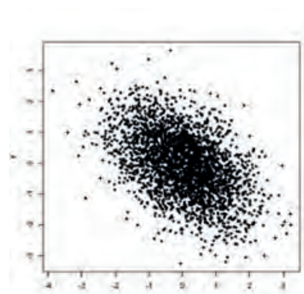


(ζ)

**B)** Για δύο ποσοτικές μεταβλητές X και Y να αντιστοιχίσετε τους συντελεστές γραμμικής συσχέτισης $r_1=0{,}8$, $r_2=0{,}2$ και $r_3= -0{,}5$, με τα αντίστοιχα διαγράμματα διασποράς.


(α)


(β)


(γ)

**Λύση**

**Α)** Σύμφωνα με τις τιμές που παίρνει ένας συντελεστής γραμμικής συσχέτισης μεταξύ δύο ποσοτικών μεταβλητών προκύπτει ότι η τιμή του $r_1$ = -1 αντιστοιχεί στο διάγραμμα (β) , η τιμή του $r_2$ = -0,8 στο διάγραμμα (δ), η τιμή του $r_3$ = -0,2 στο διάγραμμα (στ), η τιμή του $r_4$ = 0 στο διάγραμμα (ζ), η τιμή του $r_5$ = 0,2 στο διάγραμμα (ε), η τιμή του $r_6$ = 0,8 στο διάγραμμα (γ) και η τιμή του $r_7$ = 1 στο διάγραμμα (α).

**Β)** Σε καθένα από αυτά τα διαγράμματα διασποράς βλέπουμε ένα σμήνος σημείων σε αντίθεση με τα διαγράμματα διασποράς του Α που είχαμε λίγα σημεία. Το διάγραμμα διασποράς (α) παρουσιάζει ασθενή θετική γραμμική συσχέτιση με συντελεστή γραμμικής συσχέτιση $r_2$=0,2. Το διάγραμμα διασποράς (β) παρουσιάζει ισχυρή θετική γραμμική συσχέτιση με συντελεστή γραμμικής συσχέτισης $r_1$=0,8. Τέλος, το διάγραμμα διασποράς (γ) παρουσιάζει αρνητική γραμμική συσχέτιση με συντελεστή γραμμικής συσχέτισης $r_3$ = -0,5.

## Εφαρμογή 2

Στον Πίνακα 2 δίνονται τα προσδόκιμα ζωής γυναικών (Π.Ζ.Γ.) και ανδρών (Π.Ζ.Α.) του έτους 2005 για 30 τυχαία επιλεγμένες χώρες.

| Χώρα | Π.Ζ.Γ. | Π.Ζ.Α. | Χώρα | Π.Ζ.Γ. | Π.Ζ.Α. |
|---|---|---|---|---|---|
| Αίγυπτος | 63 | 60 | Λευκορωσία | 76 | 66 |
| Αυστρία | 79 | 73 | Λιθουανία | 77 | 68 |
| Αφγανιστάν | 44 | 45 | Μαλαισία | 72 | 66 |
| Βέλγιο | 79 | 73 | Μποτσουάνα | 66 | 60 |
| Βολιβία | 64 | 59 | Νησιά Μπαρμπάντος | 78 | 73 |
| Δομινικανή Δημοκρατία | 70 | 66 | Νιγηρία | 57 | 54 |
| Ελ Σαλβαδόρ | 69 | 64 | Νικαράγουα | 67 | 61 |
| Ελλάδα | 80 | 75 | Ολλανδία | 81 | 75 |
| Ζάμπια | 45 | 44 | Περού | 67 | 63 |
| Ινδία | 59 | 58 | Ρωσία | 74 | 64 |
| Ισημερινός | 73 | 67 | Σενεγάλη | 58 | 55 |
| Καμερούν | 58 | 55 | Σομαλία | 55 | 54 |
| Καναδάς | 81 | 74 | Τανζανία | 45 | 41 |
| Κίνα | 69 | 67 | Τσεχία | 77 | 69 |
| Κουβέιτ | 78 | 73 | Χιλή | 78 | 71 |

*Πίνακας 2: Προσδόκιμα ζωής των γυναικών (Π.Ζ.Γ.) και ανδρών (Π.Ζ.Α.) για 30 χώρες*

**α)** Να κατασκευάσετε το διάγραμμα διασποράς μεταξύ των προσδόκιμων ζωής γυναικών και ανδρών και να περιγράψετε το είδος της σχέσης που φαίνεται να έχουν οι δύο μεταβλητές.

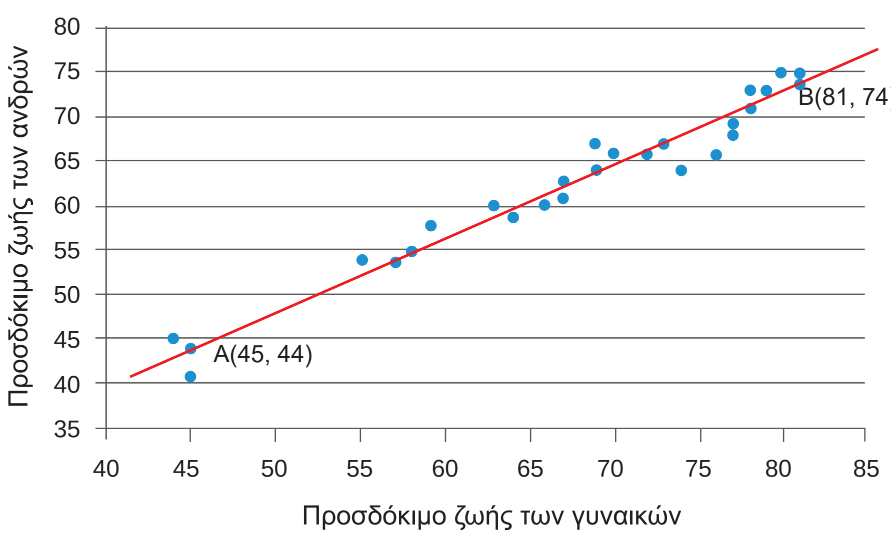
**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των προσδόκιμων ζωής γυναικών και ανδρών.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Αν για την Αλβανία, το προσδόκιμο ζωής των γυναικών είναι τα 75 έτη, μπορείτε να εκτιμήσετε το προσδόκιμο ζωής των ανδρών;

**<u>Λύση</u>**

**α)** Θεωρούμε ως ανεξάρτητη μεταβλητή (Χ) το προσδόκιμο ζωής των γυναικών και ως εξαρτημένη μεταβλητή (Υ) το προσδόκιμο ζωής των ανδρών για 30 χώρες. Βέβαια, στην περίπτωσή μας είναι τέτοια η φύση των μεταβλητών που θα μπορούσε να είναι και ανάποδα.




*Σχήμα 2: Διάγραμμα Διασποράς Π.Ζ.Γ και Π.Ζ.Α. για 30 χώρες και ευθεία προσαρμοσμένη «με το μάτι» για τα δεδομένα του πίνακα 2*

Από το διάγραμμα διασποράς φαίνεται ότι τα σημεία του διαγράμματος είναι κοντά σε μια νοητή ευθεία. Παρατηρούμε ότι, καθώς αυξάνεται το προσδόκιμο ζωής των γυναικών, φαίνεται να αυξάνεται και το προσδόκιμο ζωής των ανδρών.

**β)** Ο συντελεστής γραμμικής συσχέτισης υπολογίζεται από τη σχέση (1). Επομένως, υπολογίζουμε πρώτα τη μέση τιμή και την τυπική απόκλιση για κάθε ποσοτική μεταβλητή: $\overline{x} = 67{,}97, \overline{y} = 63{,}1, s_X = 10{,}93, s_Y = 9{,}16$ και το άθροισμα $\sum_{i=1}^{30} x_i y_i = 131611$. Οπότε, r=0,98. Αυτό σημαίνει ότι το προσδόκιμο ζωής των γυναικών και των ανδρών στο δείγμα μας είναι ισχυρά θετικά γραμμικά συσχετισμένα.

**γ)** Τα σημεία $(x_i, y_i)$,i = 1, 2,..., 30 είναι συγκεντρωμένα γύρω από μια ευθεία, όπως φαίνεται στο σχήμα 2. Η εξίσωση της ευθείας δίνεται από τη σχέση (2). Επιλέγουμε δύο σημεία, έστω τα Α(45,44) και Β(81,74) πάνω στην ευθεία που φέραμε «με το μάτι». Αντικαθιστούμε τις συντεταγμένες (x,y) των σημείων αυτών στην (2), οπότε προκύπτει το σύστημα:

$$\begin{cases} y_A = \alpha + \beta x_A \\ y_B = \alpha + \beta x_B \end{cases} \Leftrightarrow \begin{cases} 44 = \alpha + 45\beta \\ 74 = \alpha + 81\beta \end{cases}$$

- Επιλύοντας το σύστημα αυτό βρίσκουμε α = -173,28 και β = 1,39, οπότε η εξίσωση της ευθείας (1) γίνεται:

$$y = 6{,}5+0{,}83x$$

Επομένως, η ευθεία που κατά τη γνώμη μας προσαρμόζεται καλύτερα στα σημεία του διαγράμματος διασποράς διέρχεται από το σημείο (0, 6,5) και έχει συντελεστή διεύθυνσης 0,83.

**δ)** Αν για την Αλβανία, το προσδόκιμο ζωής των γυναικών είναι τα 75 έτη, χρησιμοποιώντας την παραπάνω ευθεία μπορούμε να εκτιμήσουμε το προσδόκιμο ζωής των ανδρών της Αλβανίας. Έτσι, $y=6{,}5+0{,}83\cdot 75=68{,}75$, δηλαδή 69 περίπου έτη.

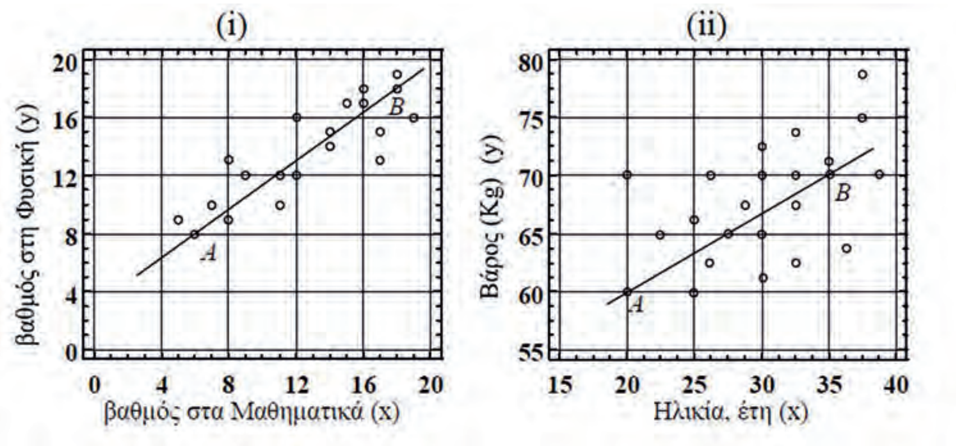
Η θετική συσχέτιση μεταξύ Π.Ζ.Γ και Π.Ζ.Α. δεν μπορεί να αποδοθεί σε αιτιακή σχέση μεταξύ των Π.Ζ.Γ και Π.Ζ.Α., αλλά στην ύπαρξη συγχυτικών παραγόντων όπως, το επίπεδο διαβίωσης, το ΑΕΠ, την ποιότητα των υπηρεσιών υγείας κάθε χώρας, κτλ., που επηρεάζουν και τα δύο προσδόκιμα ζωής.

## Ασκήσεις – Προβλήματα – Δραστηριότητες

**1)** Να διατάξετε τις παρακάτω τιμές του r σε αύξουσα τάξη του βαθμού γραμμικής συσχέτισης δύο ποσοτικών μεταβλητών X και Y: -0,6, 0,9, -0,7, 0,2, 0, -1.

**2)** Ο συντελεστής γραμμικής συσχέτισης δύο μεταβλητών X και Y είναι 0,96, ενώ ο συντελεστής γραμμικής συσχέτισης δύο μεταβλητών Ζ και Φ είναι -0,96. Ποια είναι η διαφορά τους;

**3)** Δίνονται δυο διαγράμματα διασποράς με χαραγμένες «με το μάτι» δύο ευθείες από έναν μαθητή.




**α)** Χρησιμοποιώντας τα σημεία Α και Β να βρείτε τις εξισώσεις των δύο ευθειών.

**β)** Πώς θα μπορούσατε να χρησιμοποιήσετε τις ευθείες του ερωτήματος α);

**4)** Μια εταιρεία διαφημίσεων παρουσίασε τον επόμενο πίνακα:

| Αριθμός διαφημίσεων | Έσοδα από πωλήσεις |
|---|---|
| 10 | 20000 |
| 18 | 28000 |
| 24 | 35000 |
| 32 | 44000 |
| 35 | 48000 |
| 37 | 50000 |
| 42 | 55000 |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς και να εκτιμήσετε από αυτό τον συντελεστή γραμμικής συσχέτισης μεταξύ του αριθμού διαφημίσεων της εταιρείας και των εσόδων της από τις πωλήσεις.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ του αριθμού διαφημίσεων της εταιρείας και των εσόδων της από τις πωλήσεις.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Αν το πλήθος των διαφημίσεων που αγόραζε ένα κανάλι ήταν 30, πόσο εκτιμάτε ότι θα ήταν τα έσοδα της εταιρείας διαφημίσεων;

**ε)** Αν το πλήθος των διαφημίσεων που αγόραζε ένα κανάλι ήταν 60, θα ήταν ασφαλές να εκτιμήσετε τα έσοδα της εταιρείας διαφημίσεων, όπως στο δ);

**5)** Τα δεδομένα του επόμενου πίνακα παριστάνουν τους βαθμούς (στην κλίμακα του 100) 10 μαθητών/τριών της Β΄ τάξης του Γενικού Λυκείου στα μαθήματα της Φυσικής (Χ) και των Μαθηματικών (Υ) κορμού σε μια γραπτή αξιολόγηση.

| Βαθμός- Χ | Βαθμός- Υ | Βαθμός- Χ | Βαθμός- Υ |
|---|---|---|---|
| 67 | 63 | 81 | 85 |
| 74 | 67 | 93 | 89 |
| 67 | 70 | 81 | 89 |
| 78 | 74 | 96 | 96 |
| 89 | 81 | 89 | 100 |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς και να εκτιμήσετε από αυτό τον συντελεστή γραμμικής συσχέτισης μεταξύ των βαθμών της Φυσικής (Χ) και των Μαθηματικών (Υ) των 10 μαθητών/τριών της Β΄ τάξης του Γενικού Λυκείου.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των βαθμών της Φυσικής (Χ) και των Μαθηματικών (Υ) των 10 μαθητών/τριών της Β΄ τάξης του Γενικού Λυκείου.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Πώς θα μπορούσατε να εκτιμήσετε τον βαθμό των Μαθηματικών ενός μαθητή της Β΄ Λυκείου, εάν γνωρίζατε ότι στη Φυσική έγραψε 70;

**6)** Στον παρακάτω πίνακα δίνονται οι ηλικίες και οι (συστολικές) πιέσεις αίματος 10 γυναικών.

| Ηλικία (x) | 56 | 42 | 72 | 36 | 63 | 47 | 55 | 49 | 38 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|
| Πίεση αίματος (y) | 17 | 12 | 14 | 10 | 13 | 9 | 11 | 8 | 11 | 15 |

**α)** Να σημειώσετε σε ορθοκανονικό σύστημα αξόνων τα σημεία (x, y) σύμφωνα με τον παραπάνω πίνακα, όπου x είναι η ηλικία των γυναικών σε έτη και y είναι η πίεση αίματος των γυναικών σε cm Hg.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των μεταβλητών της ηλικίας των γυναικών σε έτη (x) και της πίεσης τους σε cm Hg (y).

**7)** Στον ακόλουθο πίνακα παρουσιάζονται στην πρώτη γραμμή οι τιμές (σε €) για διαφορετικά κράνη ποδηλασίας και στη δεύτερη γραμμή η βαθμολογία ποιότητάς τους που έγινε από ειδικούς (σε μια κλίμακα από 0 έως 100, όπου όσο μεγαλύτερη είναι η τιμή τόσο πιο ποιοτικό είναι το κράνος).

| **Τιμή (€)** | **35** | **22** | **33** | **42** | **50** | **23** | **29** | **18** | **39** | **28** | **20** | **25** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Βαθμολογία ποιότητας** | **64** | **60** | **58** | **55** | **54** | **45** | **47** | **43** | **42** | **41** | **40** | **32** |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς.

**β)** Υπάρχει γραμμική συσχέτιση ανάμεσα στην τιμή και τη βαθμολογία ποιότητας;

**γ)** Θα μπορούσαμε να πούμε με βεβαιότητα ότι αν αγοράσουμε πιο φθηνό κράνος θα έχει πιο χαμηλή ποιότητα;

**δ)** Να σχεδιάσετε «με το μάτι» στο διάγραμμα διασποράς μια ευθεία που θα μπορούσε να περιγράφει τη σχέση του αναμενόμενου βαθμού ποιότητας ενός ποδηλατικού κράνους με την τιμή του.

**8)** Στον ακόλουθο πίνακα παρουσιάζονται οι συντελεστές γραμμικής συσχέτισης των γραπτών βαθμολογιών στις εξετάσεις Ιουνίου σε 5 μαθήματα ενός τμήματος Β΄ τάξης γενικού λυκείου.

| | **Άλγεβρα** | **Βιολογία** | **Γλώσσα** | **Φυσική** | **Χημεία** |
|---|---|---|---|---|---|
| **Άλγεβρα** | **1,00** | | | | |
| **Βιολογία** | **0,54** | **1,00** | | | |
| **Γλώσσα** | **0,76** | **0,81** | **1,00** | | |
| **Φυσική** | **0,70** | **0,73** | **0,71** | **1,00** | |
| **Χημεία** | **0,41** | **0,80** | **0,67** | **0,66** | **1,00** |

Να εξετάσετε κατά πόσο υπάρχει, ισχυρή ή όχι, γραμμική συσχέτιση ανάμεσα στις βαθμολογίες στα 5 εξεταζόμενα μαθήματα των μαθητών/τριών αυτών.

**9)** Το πλήθος x των οχημάτων σε εκατομμύρια και ο αριθμός y των ατυχημάτων σε εκατοντάδες, σε 15 διαφορετικές χώρες, δίνονται από τον επόμενο πίνακα:

| **Χώρα** | Α | Β | Γ | Δ | Ε | Ζ | Η | Θ | Ι | Κ | Λ | Μ | Ν | Ξ | Ο |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **x** | 8,6 | 13,4 | 12,8 | 9,3 | 1,3 | 9,4 | 13,1 | 4,9 | 13,5 | 9,6 | 7,5 | 9,8 | 23,3 | 21 | 19,4 |
| **y** | 33 | 51 | 30 | 48 | 12 | 23 | 46 | 18 | 36 | 50 | 34 | 35 | 95 | 99 | 69 |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς μεταξύ των μεταβλητών του πλήθους των οχημάτων και του αριθμού των οχημάτων για τις 15 χώρες.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των μεταβλητών του πλήθους των οχημάτων και του αριθμού των οχημάτων για τις 15 χώρες.

**10)** Από 8 γάμους που έγιναν σε μια εκκλησία ενός χωριού κατά τη διάρκεια ενός μηνός οι ηλικίες των ανδρόγυνων ήταν:

| Ηλικία γαμπρού | 20 | 22 | 24 | 25 | 28 | 30 | 33 | 38 |
|---|---|---|---|---|---|---|---|---|
| Ηλικία νύφης | 20 | 20 | 22 | 27 | 24 | 25 | 28 | 34 |

**α)** Να κατασκευάσετε το διάγραμμα διασποράς μεταξύ των ηλικιών της νύφης (Y) και του γαμπρού (X) και να περιγράψετε το είδος της σχέσης που φαίνεται να έχουν οι δύο μεταβλητές.

**β)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των ηλικιών νύφης και γαμπρού.

**γ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**δ)** Να βρείτε την αναμενόμενη ηλικία της νύφης για έναν υποψήφιο γαμπρό ετών 34.

**11)** Δίνεται δείγμα ν ζευγών παρατηρήσεων $(x_i , y_i)$ , i=1,2,…,ν δύο μεταβλητών X και Y και έστω r, ο συντελεστής γραμμικής συσχέτισης των μεταβλητών X και Y. Να αποδείξετε ότι αν όλα τα παραπάνω σημεία βρίσκονται πάνω σε μια ευθεία με εξίσωση:

$$y_i = \alpha + \beta x_i \ , \ i = 1,2,\ldots,\nu$$

τότε:

$$r = 1 \text{ αν } \alpha > 0 \text{ και } r = -1 \text{ αν } \alpha < 0.$$

# Πρόσθετο Υλικό

## Θέματα για διερεύνηση

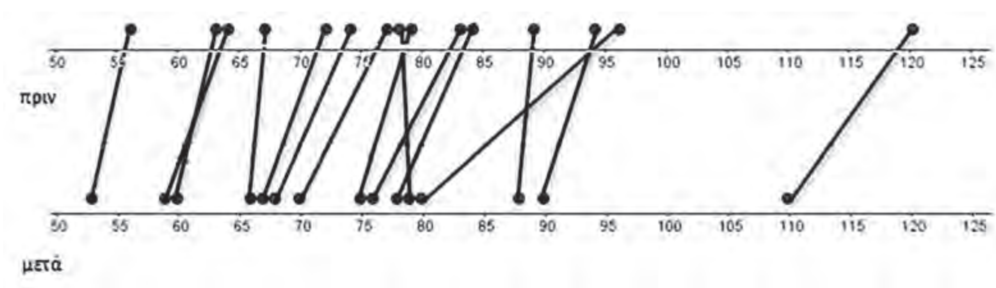

**1)** Σε μια έρευνα που έγινε με σκοπό να διερευνηθεί η αποτελεσματικότητα μιας δίαιτας, μετρήθηκε το βάρος 15 ατόμων πριν και μετά τη δίαιτα.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Πριν | 89 | 83 | 78 | 56 | 96 | 120 | 67 | 63 | 72 | 74 | 79 | 94 | 84 | 64 | 77 |
| Μετά | 88 | 76 | 79 | 53 | 80 | 110 | 66 | 60 | 67 | 67 | 75 | 90 | 78 | 59 | 70 |

**α)** Να υπολογίσετε τη μέση τιμή και την τυπική απόκλιση της μεταβλητής του βάρους των 15 ατόμων πριν ($X_{πριν}$) και μετά ($X_{μετά}$) τη δίαιτα. Να σχολιάσετε τα αποτελέσματα.

**β)** Να κατασκευάσετε τα θηκογράμματα για τη μεταβλητή του βάρους των 15 ατόμων πριν και μετά τη δίαιτα. Να τα συγκρίνετε και να διατυπώσετε την άποψή σας για το αν υπάρχει διαφορά πριν και μετά και τη δίαιτα.

**γ)** Δίνονται τα επόμενα σημειογράμματα για τη μεταβλητή του βάρους των 15 ατόμων πριν και μετά τη δίαιτα, στα οποία έχουν ενωθεί τα βάρη των 15 ατόμων πριν και μετά τη δίαιτα. Τι παρατηρείτε από το γράφημα; Επιβεβαιώνεται η παρατήρησή σας, συγκρίνοντας με τα αποτελέσματα των προηγούμενων ερωτημάτων;




**δ)** Να δημιουργήσετε τη μεταβλητή $Z = X_{πριν} - X_{μετά}$ για κάθε άτομο και να κατασκευάσετε το θηκόγραμμά της. Πώς μπορείτε να αναδείξετε τη διαφορά των τιμών του βάρους των 15 ατόμων πριν και μετά τη δίαιτα από αυτό το θηκόγραμμα;

**ε)** Να κατασκευάσετε τα διαγράμματα διασποράς ανάμεσα στις μεταβλητές του βάρους ($X_{πριν}$) και ($X_{μετά}$) των 15 ατόμων.

**στ)** Να υπολογίσετε και να ερμηνεύσετε τον συντελεστή γραμμικής συσχέτισης μεταξύ των μεταβλητών του βάρους ($X_{πριν}$) και ($X_{μετά}$) των 15 ατόμων.

**ζ)** Να σχεδιάσετε «με το μάτι» την ευθεία που φαίνεται να προσαρμόζεται καλύτερα στα δεδομένα.

**η)** Θα μπορούσατε να εκτιμήσετε το βάρος ενός ατόμου που πρόκειται να ακολουθήσει αυτή τη δίαιτα, εάν το αρχικό του βάρος ήταν 91 κιλά;

**2)** Επισκεφθείτε την ιστοσελίδα της ΕΛ.ΣΤΑΤ. (www.statistics.gr) και επιλέξτε ένα κατάλληλο σύνολο δεδομένων που να περιέχει τουλάχιστον δύο ποσοτικές μεταβλητές. Αποφασίστε ποια θα είναι η εξαρτημένη και ποια η ανεξάρτητη μεταβλητή. Στη συνέχεια, κάντε το διάγραμμα διασποράς των δύο μεταβλητών, υπολογίστε και ερμηνεύστε τον συντελεστή γραμμικής συσχέτισης και χαράξτε με το μάτι μια ευθεία που φαίνεται να προσαρμόζεται καλά στα δεδομένα σας. Τέλος, κάντε μια-δυο εκτιμήσεις για διάφορες τιμές της ανεξάρτητής σας μεταβλητής που να ανήκουν στο εύρος τιμών της. Καταγράψτε τα αποτελέσματα σας, συντάσσοντας μια μικρή αναφορά (μέχρι 2 σελίδες, με πίνακες και γραφήματα).

**3)** Να ετοιμάσετε ένα ερωτηματολόγιο για να μελετήσετε ορισμένα χαρακτηριστικά από την καθημερινότητά σας και να τα συσχετίσετε με τον γενικό βαθμό που πήραν κατά το τελευταίο τετράμηνο. Τα χαρακτηριστικά που μπορείτε να καταγράψετε είναι: πόσες ώρες διαβάζετε καθημερινά κατά μέσο όρο, πόσες εξωσχολικές δραστηριότητες έχετε, πόσες ώρες κατά μέσο όρο ξοδεύετε σε αυτές και ό,τι άλλο θα επιθυμούσατε εσείς. Αναλύστε τα δεδομένα με αυτά που μάθατε σε αυτήν την παράγραφο. Περιγράψτε τη σχέση που έχει καθένα από αυτά τα χαρακτηριστικά με τον βαθμό του τετραμήνου. Καταγράψτε τα αποτελέσματά σας, συντάσσοντας μια μικρή αναφορά (μέχρι 5 σελίδες, με πίνακες και γραφήματα).

Κωδικός Βιβλίου: 0-22-0269
ISBN 978-960-06-6144-6

ΙΤΥΕ
"ΔΙΟΦΑΝΤΟΣ"
ινστιτουτο
τεχνολογιας
υπολογιστων & εκδοσεων

(01) 000000 0 22 0269 3