

ΣΤΑΤΙΣΤΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΕΛΕΓΧΟΣ ΠΟΙΟΤΗΤΑΣ ΠΕΙΡΑΜΑΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΣΤΗ ΧΗΜΙΚΗ ΑΝΑΛΥΣΗ

"The greatest enemy of knowledge is not ignorance; it is the illusion of knowledge"

Stephen Hawking

"Δεν υπάρχει λόγος να είσαι τόσο ακριβής, όταν δεν γνωρίζεις καν για ποιο πράγμα μιλάς"

John von Neumann, Ούγγρος Μαθηματικός

Η επιστήμη εξαρτάται από πρωτότυπες σκέψεις. Μία είναι όταν κάνουμε μια υπόθεση (hypothesis). Μια άλλη είναι όταν επινοούμε μια δοκιμή (δοκιμασία) ή ένα πείραμα για να δείξουμε πόσο πιθανή είναι η υπόθεση που κάναμε.

Περιεχόμενα του μαθήματος

1. ΒΑΣΙΚΗ ΟΡΟΛΟΓΙΑ ΣΤΗΝ ΑΝΑΛΥΤΙΚΗ ΧΗΜΕΙΑ
2. ΠΕΡΙΓΡΑΦΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ
3. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΣΤΗ ΘΕΩΡΙΑ ΠΙΘΑΝΟΤΗΤΩΝ
4. ΕΚΤΙΜΗΣΗ ΣΧΕΤΙΚΩΝ ΠΑΡΑΜΕΤΡΩΝ ΠΛΗΘΥΣΜΟΥ - ΔΟΚΙΜΕΣ ΣΗΜΑΝΤΙΚΟΤΗΤΑΣ
5. ΓΡΑΦΗΜΑΤΑ ΕΛΕΓΧΟΥ ΠΟΙΟΤΗΤΑΣ
6. ΜΕΘΟΔΟΙ ΒΑΘΜΟΝΟΜΗΣΗΣ: ΣΥΜΜΕΤΑΒΟΛΗ ΚΑΙ ΣΥΣΧΕΤΙΣΗ

Περιεχόμενα του μαθήματος

7. ΔΙΕΡΓΑΣΤΗΡΙΑΚΕΣ ΔΟΚΙΜΕΣ
8. ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ (ΑΝΟΝΑ)
9. ΕΚΤΙΜΗΣΗ ΑΒΕΒΑΙΟΤΗΤΑΣ ΣΤΗ ΧΗΜΙΚΗ ΑΝΑΛΥΣΗ
10. ΕΙΣΑΓΩΓΗ ΣΤΟΝ ΠΕΙΡΑΜΑΤΙΚΟ ΣΧΕΔΙΑΣΜΟ

ΒΑΣΙΚΗ ΟΡΟΛΟΓΙΑ ΣΤΗΝ ΑΝΑΛΥΤΙΚΗ ΧΗΜΕΙΑ

Οι αναλυτικοί χημικοί επικοινωνούν μεταξύ τους χρησιμοποιώντας μια ορολογία που τους επιτρέπει να αποδίδουν και να αξιολογούν κάποιες ειδικές έννοιες.

Η συζήτηση γύρω από την Αναλυτική Χημεία και η κατανόηση των αρχών της προϋποθέτει σωστή γνώση της συγκεκριμένης γλώσσας.

Πολλοί χημικοί είναι, ίσως, εξοικειωμένοι με έννοιες όπως: μήτρα δείγματος, αναλυτική μέθοδος..., είναι όμως πιθανό να μη έχει δοθεί σ' αυτές το κατάλληλο αναλυτικό-εννοιολογικό περιεχόμενο.

- Εισαγωγή κοινών όρων που καθημερινά χρησιμοποιούνται από τους αναλυτικούς χημικούς
- ταξινόμηση και ερμηνεία τους για την αποφυγή συγχύσεων

Η χημεία στηρίζεται σε τρεις συσχετιζόμενους θεμελιώδεις επιστημονικούς λίθους: τη **θεωρία**, τη **σύνθεση** και την **ανάλυση**.

Η Αναλυτική Χημεία καλύπτει - όπως είναι εύλογο - την **ανάλυση**.

Η Αναλυτική Χημεία, επιπλέον, εμπίπτει και στο πλαίσιο της **μετρολογίας** (η επιστήμη των μετρήσεων). Κάθε μέτρηση συνοδεύεται από αβεβαιότητα.

Στόχος: η **ελαχιστοποίηση της αβεβαιότητας στην πληροφορία που προκύπτει**.

Η περιγραφή του βασικού πεδίου της Αναλυτικής Χημείας που είναι η απόδοση (βιο)χημικών μετρήσεων με στόχο την παραγωγή ποιοτικών και ποσοτικών αποτελεσμάτων και τη διευκρίνιση δομών, απαιτεί τη γνώση χαρακτηριστικών της μετρολογίας.

Η Αναλυτική Χημεία, εκ φύσεως, αποτελεί επιστημονικό κλάδο που παρέχει (βιο)χημικές πληροφορίες με σκοπό τη λήψη έγκαιρων αποφάσεων σε πολλά κοινωνικά και οικονομικά πεδία.

Η Αναλυτική Χημεία έχει οριστεί με πολλούς τρόπους:

Ομοσπονδία Ευρωπαϊκών Χημικών Κοινοτήτων (Federation of European Chemical Societies - FECS): Αναλυτική Χημεία είναι ένας επιστημονικός κλάδος που αναπτύσσει και εφαρμόζει μεθόδους, οργανολογίες και στρατηγικές για την απόκτηση πληροφοριών σχετικά με τη σύσταση και τη φύση της ύλης στο χώρο και στο χρόνο.

IUPAC 2002, Analytical Chemistry Division: *Analytical chemistry is a scientific discipline that develops and applies methods, instruments and strategies to obtain information on the composition and nature of matter in space and time, as well as on the value of these measurements, i.e., their uncertainty, validation, and/or traceability to fundamental standards.*

Συμπληρωματικά, ένας ορισμός που κατ' ουσία δεν απέχει πολύ από τον προηγούμενο: Αναλυτική Χημεία είναι ένας επιστημονικός κλάδος της μετρολογίας που αναπτύσσει, βελτιστοποιεί και εφαρμόζει διεργασίες μέτρησης με σκοπό τη λήψη (βιο)χημικών ποιοτικών πληροφοριών συνολικής και μερικής μορφής από φυσικά και τεχνητά αντικείμενα ή συστήματα για να λύσει αναλυτικά προβλήματα.

Τελικά, βασικό αντικείμενο της Αναλυτικής Χημείας αποτελεί η απόκτηση όσο το δυνατόν περισσότερων (βιο)χημικών πληροφοριών της μέγιστης δυνατής ποιότητας, για αντικείμενα και συστήματα, με την ελάχιστη κατανάλωση δείγματος, χρόνου, ανθρώπινου δυναμικού, κόστους και κινδύνου.

Στο βαθμό, λοιπόν, που η Αναλυτική Χημεία αποτελεί επιστημονικό κλάδο (βιο)χημικών πληροφοριών πρώτος στόχος είναι η **ελαχιστοποίηση της αβεβαιότητας στην πληροφορία που προκύπτει.**

Ανάλυση: η εφαρμογή μιας ή περισσότερων αναλυτικών μεθόδων για τη λήψη χημικών ή φυσικών πληροφοριών γύρω από το δείγμα. Ο όρος ανάλυση αναφέρεται μόνο στο δείγμα και ποτέ στις ουσίες που αυτό περιέχει.

Προσδιορισμός: το εργαλείο της ανάλυσης κατά τον οποίο μετρώνται κάποιες από τις φυσικές ή χημικές ιδιότητες του ή των συστατικών του δείγματος. Ο προσδιορισμός μπορεί να είναι ποιοτικός (ταυτοποίηση), ποσοτικός (ποσοτικοποίηση) ή έλεγχος ορίου (limit test).

Τα συστατικά δεν αναλύονται αλλά προσδιορίζονται. Το συστατικό-στόχος του προσδιορισμού λέγεται **προσδιοριζόμενο** συστατικό. Από τον αγγλικό όρο **analyte** έχει επικρατήσει και στα Ελληνικά ο όρος **αναλύτης** (σε αντίθεση με τον **αναλυτή-analyst** που είναι ο αναλυτικός επιστήμονας και τον **αναλυτή-analyzer** που είναι ένα αυτοματοποιημένο αναλυτικό όργανο που εκτελεί αναλύσεις δειγμάτων).

Μήτρα ή υπόστρωμα (matrix): αποτελεί το υπόλοιπο του δείγματος πέραν του αναλύτη

Μέτρηση (*run*): αφορά την απλή ανάγνωση/λήψη αναλυτικού σήματος που προκαλείται στο αναλυτικό σύστημα λόγω της παρουσίας του αναλύτη.

Πρωτογενή δεδομένα (*primary data*): Το σύνολο των μετρήσεων στην πιο στοιχειώδη μορφή τους

Replicates: Οι επαναλαμβανόμενες μετρήσεις ή προσδιορισμοί.

Οι προαναφερθέντες όροι σχετίζονται άμεσα με τους όρους: **αναλυτικό αποτέλεσμα** (*analytical result*), **αναλυτική αναφορά** (*analytical report*) και **αναλυτική ποιότητα** (*analytical quality*).

Αναλυτικό αποτέλεσμα: αφορά ποιοτικά και ποσοτικά δεδομένα που λαμβάνονται από μαθηματική επεξεργασία των πρωτογενών δεδομένων μετά την εκτέλεση μιας διαδικασίας μέτρησης.

Αναλυτική αναφορά: το σύνολο των αναλυτικών αποτελεσμάτων και η ερμηνεία τους σε σχέση με το αναλυτικό πρόβλημα.

Αναλυτική ποιότητα: αναφέρεται στην επίδοση της πληροφορίας που παρέχεται με σκοπό την επίλυση του αναλυτικού προβλήματος.

Περιλαμβάνει τέσσερα κύρια συστατικά: την ποιότητα των αποτελεσμάτων, την ακολουθούμενη πορεία, τα αναλυτικά εργαλεία που χρησιμοποιήθηκαν και την οργάνωση της αναλυτικής δουλειάς.

Ας υποθεθεί ότι ζητείται η εύρεση ενός τρόπου προσδιορισμού της συγκέντρωσης ενός μετάλλου σε νερό. Ποια είναι η μεθοδολογική προσέγγιση ?

Διαφοροποίηση των τεσσάρων επιπέδων αναλυτικής μεθοδολογίας: **τεχνική, μέθοδος, πορεία, πρωτόκολλο**.

Αναλυτική τεχνική: είναι η εφαρμογή μιας φυσικοχημικής αρχής (απορρόφηση, εκπομπή ακτινοβολίας, ανάπτυξη διαφοράς δυναμικού κλπ) στην ανάλυση. Οι αναλυτικές τεχνικές χωρίζονται σε δύο μεγάλες κατηγορίες ανάλογα με το εάν το αναλυτικό σήμα είναι ανάλογο της απόλυτης ή σχετικής ποσότητας του αναλύτη στο δείγμα.

Τεχνικές συνολικής ανάλυσης (*total analysis techniques*)

λίγο πολύ συμπίπτουν με τις κλασικές τεχνικές, όπως η ογκομέτρηση, η σταθμική ανάλυση αλλά και η κουλομετρία

Τεχνικές συγκέντρωσης (*concentration techniques*) με αποκλειστικό εκπρόσωπο τις ενόργανες τεχνικές ανάλυσης.

Αναλυτική μέθοδος: είναι η εφαρμογή μιας τεχνικής για τον ποιοτικό ή ποσοτικό προσδιορισμό ενός αναλύτη ή περισσοτέρων, σε μια συγκεκριμένη μήτρα. Φλογοφωτομετρική μέθοδος προσδιορισμού νατρίου, ποτενσιομετρική μέθοδος προσδιορισμού φωσφορικών ιόντων.

Χαρακτηριστικά ποιότητας:

ακρίβεια, επαναληψιμότητα, ευαισθησία, εκλεκτικότητα, ειδικότητα, αντοχή, ανθεκτικότητα, κόστος κλπ.

αποτελούν μερικά από τα κριτήρια επιλογής μεθόδου.

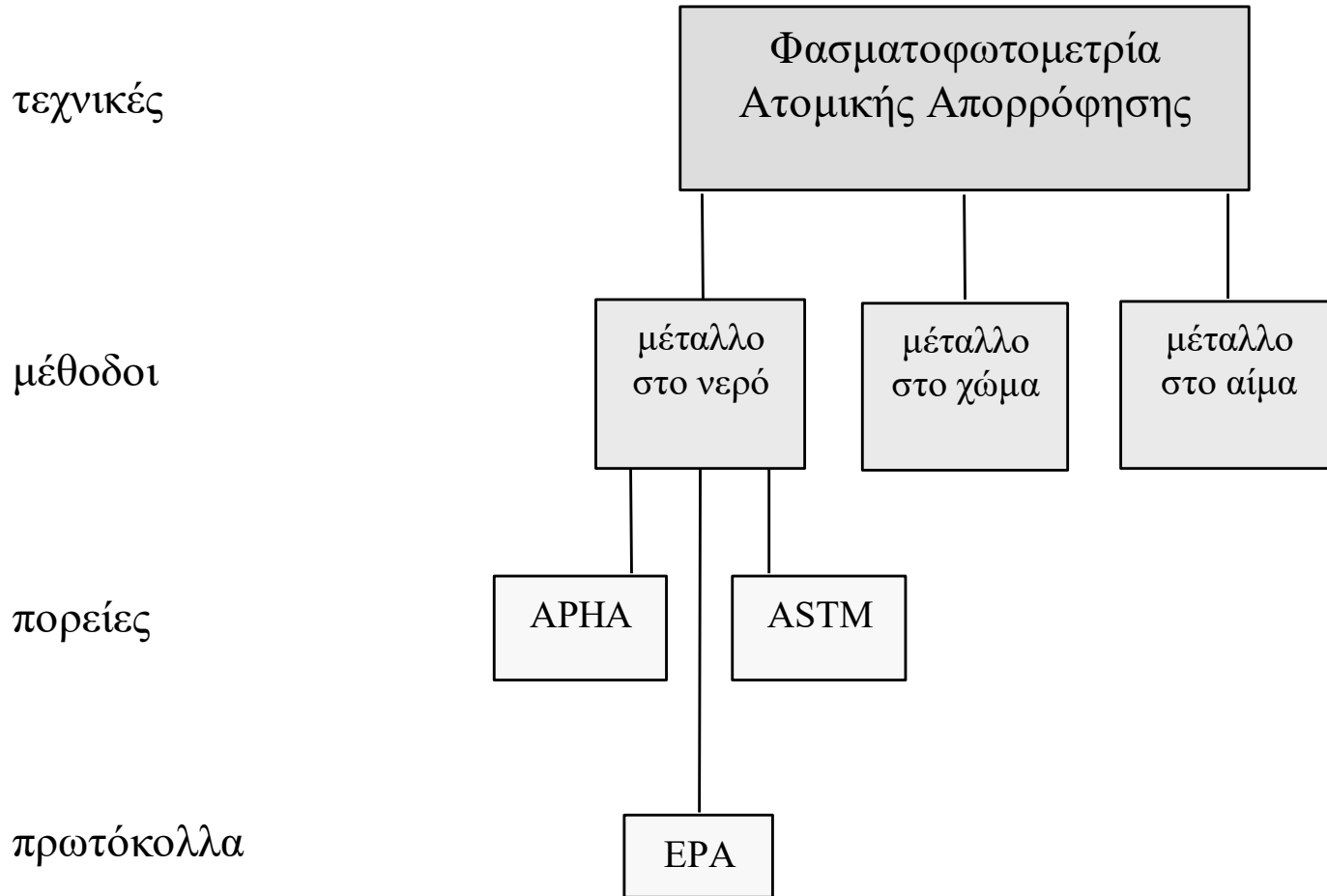
Πορεία εργασίας: η γενική περιγραφή σταδίων για την εκτέλεση της αναλυτικής μεθόδου σε ένα δείγμα και περιλαμβάνει πληροφορίες δειγματοληψίας, χειρισμού των παρεμποδίσεων και επικύρωσης των αποτελεσμάτων.

Επικύρωση (*validation*): διαδικασία που επιβεβαιώνει ότι η ακολουθούμενη πορεία παράγει αποδεκτά αποτελέσματα.

Μια αναλυτική μέθοδος δεν οδηγεί κατ' ανάγκη σε μία μόνο πορεία. Οι διαφορετικοί αναλυτές (*analysts*) και υπηρεσίες προσαρμόζουν τις μεθόδους στις κατά περίπτωση ανάγκες τους.

Πρωτόκολλο: μια σειρά αυστηρών και λεπτομερών οδηγιών συγκεκριμένης πορείας που ακολουθείται και σχετίζεται κυρίως με παγκοσμίως αποδεκτούς φορείς.

Υπάρχει μια **πολυεπίπεδη σειρά** στη μεθοδολογική προσέγγιση για τον προσδιορισμό της συγκέντρωσης μετάλλου σε δείγμα.



Διάγραμμα που απεικονίζει την ιεραρχική σχέση μεταξύ μιας τεχνικής, μιας μεθόδου, μιας πορείας και ενός πρωτοκόλλου στον προσδιορισμό μετάλλου. ΑΡΗΑ: Αμερικανική Ένωση Δημόσιας Υγείας, ASTM: Αμερικανική Κοινότητα Δοκιμής και Υλικών, EPA: Υπηρεσία Προστασίας Περιβάλλοντος ΗΠΑ.

ΒΑΘΜΟΝΟΜΗΣΗ (CALIBRATION-STANDARDIZATION)

Βαθμονόμηση συσκευής ή οργάνου: μια συσκευή ή ένα όργανο ελέγχεται για να διασφαλιστεί η σωστή λειτουργία του (αναλυτικός ζυγός, πεχάμετρο). Διατίθενται κατάλληλα πρότυπα και οδηγίες από τον κατασκευαστή.

Βαθμονόμηση μεθόδου: πειραματικός προσδιορισμός της σχέσης της απόκρισης ή του σήματος ενός οργάνου ως προς την ποσότητα του αναλύτη (standardization).

Για μια μέθοδο συνολικής ανάλυσης, η βαθμονόμηση συνήθως ορίζεται από τη στοιχειομετρία της χημικής αντίδρασης που είναι υπεύθυνη για το σήμα. Για μια μέθοδο συγκέντρωσης, απαιτείται η χάραξη μιας καμπύλης που απεικονίζει το σήμα S ως προς τη συγκέντρωση του αναλύτη στα πρότυπα διαλύματα (**καμπύλη βαθμονόμησης** ή **καμπύλη αναφοράς**).

Η κλίση της καμπύλης βαθμονόμησης είναι ένας από τους δύο παράγοντες που καθορίζουν την ευαισθησία (ο δεύτερος είναι η επαναληψιμότητα).

Ευαισθησία (sensitivity) ενός οργάνου ή μιας μεθόδου: μέτρο της ικανότητάς τους να διακρίνουν μικρές διαφορές στην ποσότητα ή συγκέντρωση του αναλύτη (A).

Η ευαισθησία k εκφράζεται, θεωρώντας ως $\Delta S(A)$ τη μικρότερη μετρήσιμη μεταβολή του σήματος και ΔnA ($\Delta C(A)$) τη μικρότερη διαφορά στην ποσότητα (συγκέντρωση) του αναλύτη που μπορεί να μετρηθεί, ως ακολούθως:

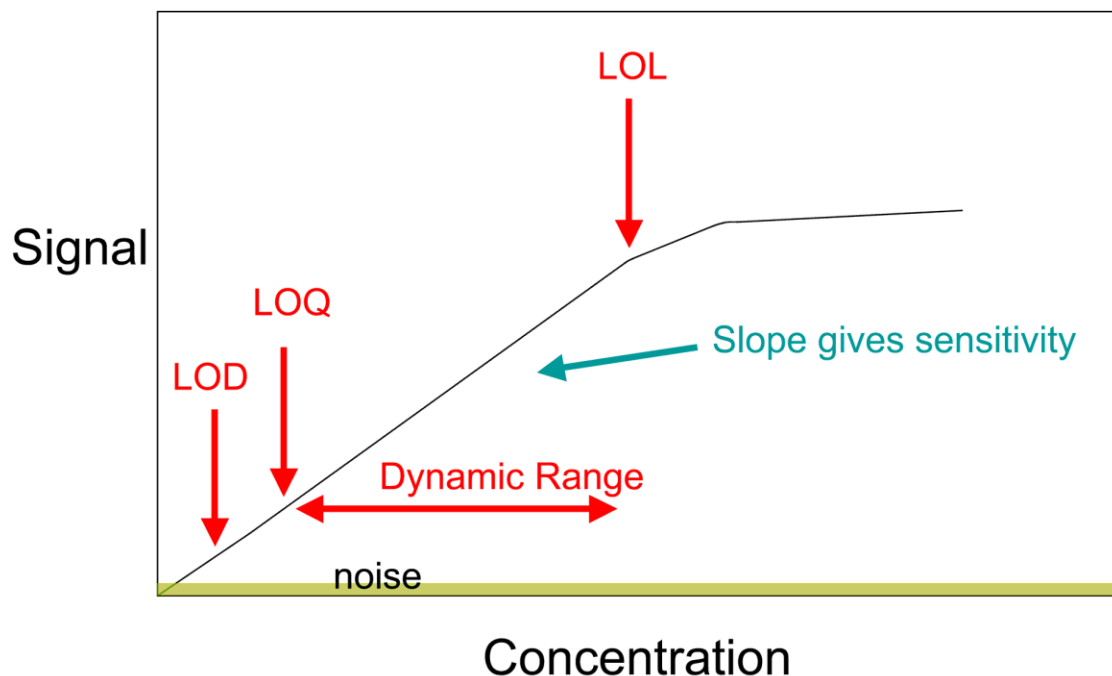
$$k = \Delta S(A) / \Delta n(A)$$

$$k = \Delta S(A) / \Delta C(A)$$

(τεχνική συνολικής ανάλυσης)

(τεχνική συγκέντρωσης)

Η ευαισθησία συχνά συγχέεται με το όριο ανίχνευσης (detection limit) της μεθόδου το οποίο είναι η μικρότερη ποσότητα ή συγκέντρωση του αναλύτη που μπορεί να ανιχνευτεί με καθορισμένη στάθμη εμπιστοσύνης.



Εκλεκτικότητα (*selectivity*) και **ειδικότητα (*specificity*)**: δύο μέτρα εκτίμησης της αξιοπιστίας μιας μέτρησης παρουσία παρεμποδιστών.

Εκλεκτικότητα: η ικανότητα της μεθόδου να διακρίνει το μετρούμενο αναλύτη από άλλες ουσίες. Παρουσία ενός παρεμποδιστή I , το αναλυτικό σήμα S περιλαμβάνει το άθροισμα των σημάτων του αναλύτη $S(A)$ του παρεμποδιστή $S(I)$ και του τυφλού $S(bI)$, ως ακολούθως:

$$S = S(A) + S(I) + S(bI) = k(A) n(A) + k(I) n(I) + S(bI) \text{ (τεχνική συνολικής ανάλυσης)}$$

$$S = S(A) + S(I) + S(bI) = k(A) C(A) + k(I) C(I) + S(bI) \text{ (τεχνική συγκέντρωσης)}$$

όπου $k(A)$ και $k(I)$ είναι οι ευαισθησίες βαθμονόμησης για τον αναλύτη και τον παρεμποδιστή αντίστοιχα και $n(A)$, $n(I)$ και $C(A)$, $C(I)$ οι αντίστοιχες ποσότητες ή συγκεντρώσεις του αναλύτη και του παρεμποδιστή.

Η εκλεκτικότητα μιας μεθόδου για τον παρεμποδιστή I σε σχέση με τον αναλύτη A ορίζεται από το **συντελεστή εκλεκτικότητας** $K_{(AI)}$.

$$K_{(AI)} = k(A) / k(I)$$

Οι συντελεστές εκλεκτικότητας αποτελούν χρήσιμα αριθμητικά μέτρα αξιολόγησης της εκλεκτικότητας αναλυτικών μεθόδων.

Ειδικότητα: γενικά, θεωρείται η πλήρης εκλεκτικότητα (100 %).

Ωστόσο, σε πολλά εγχειρίδια Αναλυτικής Χημείας δεν γίνεται διάκριση των δύο όρων ειδικότητα, εκλεκτικότητα και χρησιμοποιείται μόνον ο δεύτερος από τους δύο.

Πολλές φορές, μέτρο της ειδικότητας θεωρείται ο λόγος των συγκεντρώσεων παρεμποδιστή προς αναλύτη για τον οποίο προκύπτει σφάλμα 5% στη μέτρηση. Ωστόσο, αυτό το μέτρο ειδικότητας σχετίζεται άμεσα με τη συγκέντρωση του αναλύτη στο υπό μελέτη σύστημα.

Ο όρος **πρότυπο (standard)** έχει πολλές και σημαντικά διαφοροποιημένες έννοιες:

1. Ένα υλικό μέτρησης που χρησιμοποιείται για να συγκρίνει αναλυτικές μετρήσεις. Τέτοια είναι τα **φυσικά πρότυπα** (*physical standards*) όπως τα πρότυπα βάρη, οι **καθαρές ουσίες** ή **μίγματα** που χρησιμοποιούνται ως πρότυπα (*pure and mixed substances used as standards*) που είναι υψηλής καθαρότητας γεγονός το οποίο εξασφαλίζεται από τον κατασκευαστή και τα **πρότυπα δείγματα** (*sample standards*), όπως τα πιστοποιημένα υλικά αναφοράς (Certified Reference Materials-CRMs) που προσομοιάζουν τις ιδιότητες των πραγματικών υπό ανάλυση δειγμάτων.
2. Ένα ελάχιστο ή το **κατώφλι** (*threshold*) για τον καθορισμό των αποδεκτών χαρακτηριστικών ενός δείγματος ή ενός συστήματος (πληροί τα πρότυπα ?).
3. Ένα λεπτομερές πρωτόκολλο που θα εφαρμοστεί σε ειδικές περιπτώσεις.
4. Ένα αρχείο-έγγραφο που σχετίζεται με συστήματα ποιότητας αναλυτικών εργαστηρίων.

Τα τρία τελευταία ανήκουν στην κατηγορία των **γραπτών προτύπων** (*written standards*) και μπορεί να είναι λεπτομερείς περιγραφές αναλυτικών διεργασιών (πρότυπες μέθοδοι) (ISO, AOAC, EPA), επίσημα έγγραφα (π.χ. Νομοθεσία Ε.Ε.) που ορίζουν επιτρεπόμενες συγκεντρώσεις ουσιών σε συγκεκριμένα δείγματα ή ακόμη, μπορεί να είναι πλήρη αναλυτικά έγγραφα που περιέχουν τις αρχές και τις πρακτικές της εργαστηριακής λειτουργίας (ISO/IEC 17025:1999).

Είναι φανερό, ότι ο όρος πρότυπο περιλαμβάνει τόσο υλικά όσο και έγγραφα ή σειρά τιμών και τη σχέση μεταξύ τους.

Ορολογία που χαρακτηρίζει μία αναλυτική μέθοδο: ISO 5725 "Accuracy (Trueness and Precision) of measurement methods and results"

Accuracy (Ορθότητα, ακρίβεια...): Μέτρο της εγγύτητας της μετρούμενης τιμής προς την πραγματική (αποδεκτή) τιμή ή τιμή αναφοράς.

$E = x_i - \mu$, όπου x_i είναι η πειραματική τιμή και μ η αποδεκτή τιμή.

Precision (πιστότητα, **τυχαίο σφάλμα**...): Μέτρο της διασποράς μιας σειράς μετρήσεων από **μέτρηση σε μέτρηση** και μέσα στην **ίδια σειρά μετρήσεων** (within-run precision). Εκφράζεται με ένα στατιστικό μέτρο διασποράς (π.χ. τυπική απόκλιση ή $|x_i - \bar{x}|$).

Repeatability (επαναληψιμότητα): Μέτρο της διασποράς ομοειδών μετρήσεων που εκτελούνται στο ίδιο εργαστήριο και από το ίδιο προσωπικό.

Reproducibility (αναπαραγωγιμότητα). Αποτελεί ανάλογο μέτρο της διασποράς ομοειδών μετρήσεων, που έχουν ληφθεί σε διάφορα εργαστήρια.

Ορολογία που χαρακτηρίζει μία αναλυτική μέθοδο: ISO 5725 "Accuracy (Trueness and Precision) of measurement methods and results"

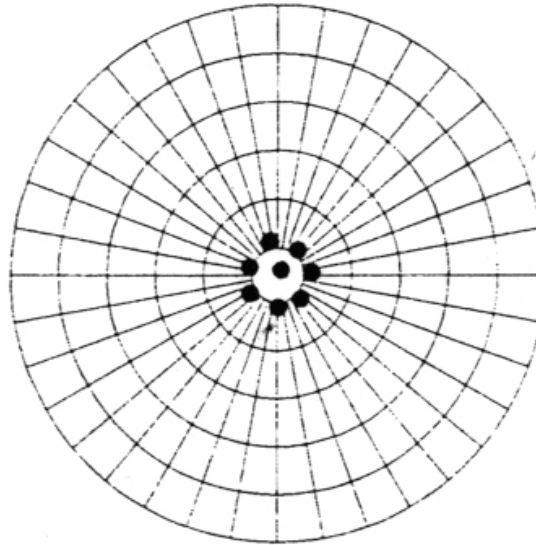
Trueness (αληθότητα): Η εγγύτητα της συμφωνίας μεταξύ της μέσης τιμής που λαμβάνεται από μια πολύ μεγάλη σειρά αποτελεσμάτων και της αποδεκτής τιμής αναφοράς ($\bar{X} - \mu$). Σχετίζεται με τη **μεροληψία**.

Bias (μεροληψία, **συστηματικό σφάλμα**...): Η διαφορά μεταξύ της αναμενόμενης (μέσης) τιμής μιας σειράς αποτελεσμάτων και της αποδεκτής τιμής αναφοράς.

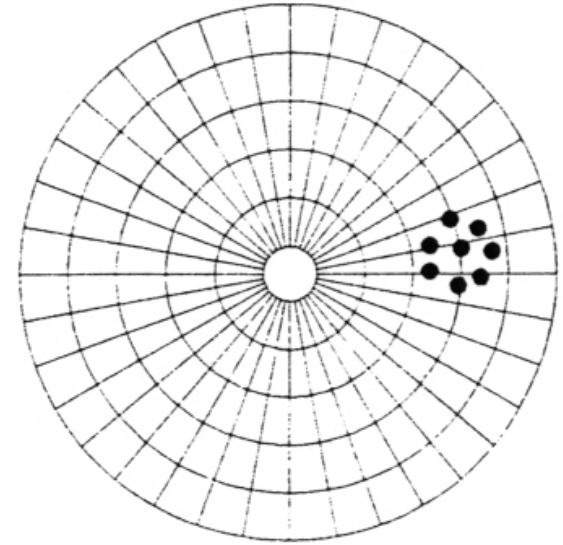
Συστηματικά και τυχαία σφάλματα

Τυχαία σφάλματα:

Επαναλαμβανόμενες μετρήσεις δεν εμπίπτουν στο ίδιο σημείο σε σχέση με τη μέση τιμή



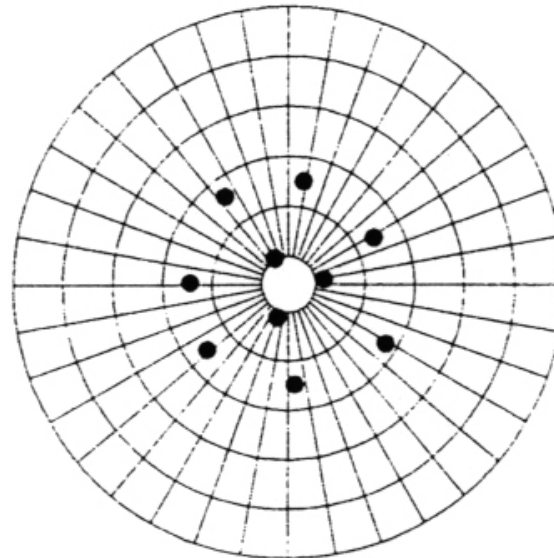
Μικρό συστηματικό σφάλμα
Μικρό τυχαίο σφάλμα



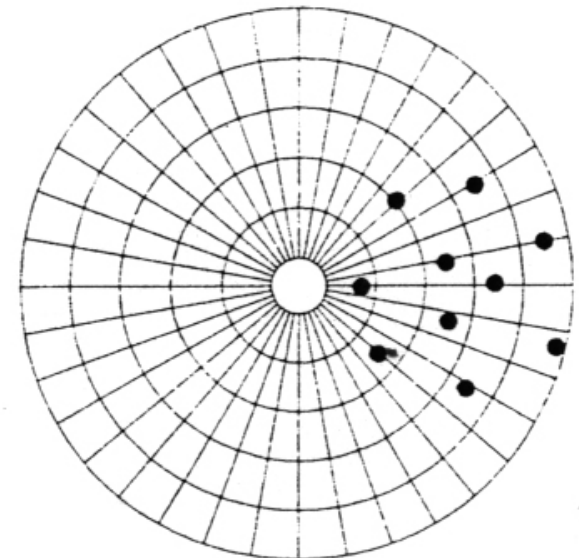
Μεγάλο συστηματικό σφάλμα
Μικρό τυχαίο σφάλμα

Συστηματικά σφάλματα:

Εμφανίζουν το σύνολο των αποτελεσμάτων ως μικρές ή μεγάλες τιμές σε σχέση με την αποδεκτή τιμή

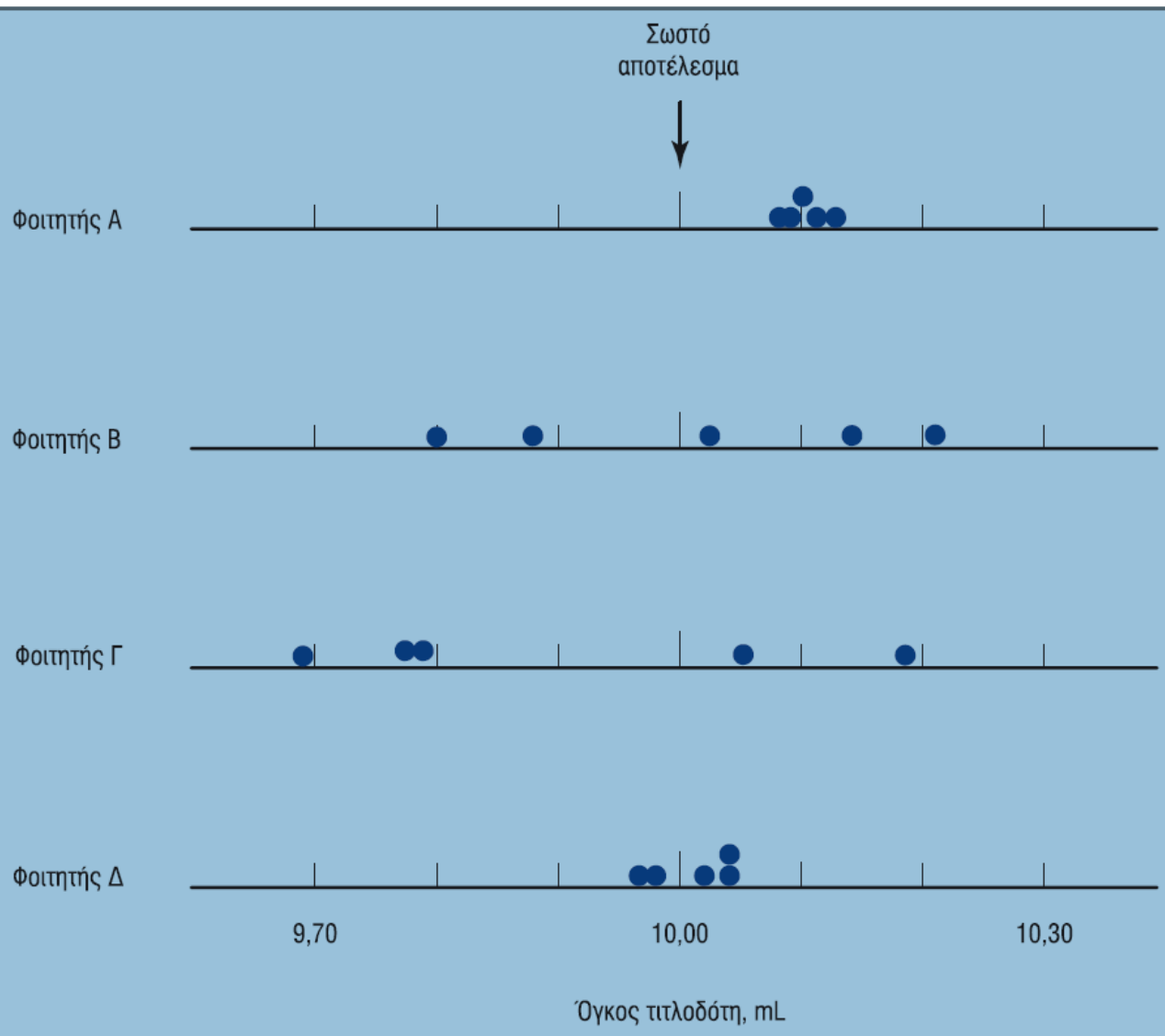


Μικρό συστηματικό σφάλμα
Μεγάλο τυχαίο σφάλμα



Μεγάλο συστηματικό σφάλμα
Μεγάλο τυχαίο σφάλμα

Συστηματικά και τυχαία σφάλματα



Τέσσερεις φοιτητές (Α-Δ) πραγματοποιούν μία ανάλυση στην οποία ακριβώς 10,00 mL NaOH συγκέντρωσης ακριβώς 0,1 M ογκομετρείται με διάλυμα HCl ακριβώς 0,1 M. Κάθε φοιτητής πραγματοποιεί πέντε επαναλήψεις.

Παρατηρήσεις επί της ορολογίας

Μια απλή μέτρηση μπορεί να θεωρηθεί ότι προκύπτει ως εξής:

$$x_i = \mu + B + e$$

(μετρούμενη τιμή) (αποδεκτή τιμή) (συστημ. σφάλμα, bias) (τυχαίο σφάλμα)

Για πολλές μετρήσεις που έχουν μια μέση τιμή $\bar{x}$ σε συνθήκες επαναληψιμότητας, οι τιμές B και e αποκτούν μέτρα και αντικαθίστανται

$$x_i = \mu + (\bar{x} - \mu) + (x_i - \bar{x})$$

(μετρούμενη τιμή) (αποδεκτή τιμή) (συστημ. σφάλμα, bias) (τυχαίο σφάλμα)

ή μετά από αναδιάταξη

$$x_i - \mu = (\bar{x} - \mu) + (x_i - \bar{x})$$

(ολικό σφάλμα μιας μέτρησης) (συστημ. σφάλμα, bias) (τυχαίο σφάλμα)

Τα παραπάνω αποτελούν μέτρα των εννοιών

$$\text{accuracy} = \text{trueness} + \text{precision}$$

Συστηματικά και τυχαία σφάλματα

Ολικό αναλυτικό σφάλμα = συστηματικό σφάλμα + τυχαίο σφάλμα

Τυχαία σφάλματα	Συστηματικά σφάλματα
Επαναλαμβανόμενες μετρήσεις δεν εμπίπτουν στο ίδιο σημείο σε σχέση με τη μέση τιμή	Εμφανίζουν το σύνολο των αποτελεσμάτων ως μικρές ή μεγάλες τιμές σε σχέση με την αποδεκτή τιμή
Επηρεάζουν την πιστότητα (precision)	Παράγουν μεροληψία (bias) δηλ. τη συνολική απόκλιση ενός αποτελέσματος από την αληθή τιμή <u>ακόμη κι όταν τα τυχαία σφάλματα είναι πολύ μικρά</u>
Μπορούν να εκτιμηθούν από επαναλαμβανόμενες μετρήσεις	Δεν είναι δυνατό να ανιχνευτούν με επαναλαμβανόμενες μετρήσεις
Μπορούν να ελαχιστοποιηθούν αλλά όχι να εξαλειφθούν	Μπορούν να διορθωθούν με σωστή βαθμονόμηση, υλικά αναφοράς (CRMs), χρήση πολλαπλών μεθόδων
Προκαλούνται από τον ανθρώπινο παράγοντα και τα όργανα	Προκαλούνται από τον ανθρώπινο παράγοντα και τα όργανα

Η στατιστική στην αναλυτική χημεία

Είναι ο κλάδος των εφαρμοσμένων μαθηματικών, ο οποίος βασίζεται σε ένα σύνολο αρχών και μεθοδολογιών με σκοπό:

1. Το σχεδιασμό της διαδικασίας συλλογής δεδομένων

Η συλλογή των δεδομένων αναφέρεται κι ως **δειγματοληψία** (sampling) κι αποτελεί ξεχωριστό πεδίο της στατιστικής.

2. Τη συνοπτική και αποτελεσματική παρουσίαση των δεδομένων

Η περιγραφή των δεδομένων και η παρουσίαση συνοπτικών στοιχείων και γραφημάτων αναφέρεται ως **περιγραφική στατιστική** (descriptive statistics).

3. Την ανάλυση και εξαγωγή συμπερασμάτων από τα δεδομένα

Η ανάλυση στατιστικών δεδομένων και η ερμηνεία των αποτελεσμάτων αναφέρεται ως **στατιστική συμπερασματολογία** (statistical inference) και αποτελεί τη βάση της στατιστικής και γι' αυτό συχνά αναφέρεται απλά και ως **στατιστική**.

Βέβαια φαινόμενα → Χρήση **καθοριστικών** μεθόδων - μοντέλων

Καθημερινό παράδειγμα

Σε πόση ώρα διανύει ένα αυτοκίνητο μια απόσταση 200 Km όταν κινείται με μέση ταχύτητα 100 Km/h?

Απάντηση . . .

Αβέβαια ή **τυχαία** φαινόμενα → Χρήση **στατιστικών** μεθόδων - μοντέλων

Παράδειγμα

Πόσο χρόνο χρειάζεται για να πραγματοποιηθεί κάποια αντίδραση?

Απάντηση

Θερμοκρασία, ταχύτητα ανάδευσης, παρουσία καταλύτη κλπ.

Το αναλυτικό πρόβλημα και η στατιστική ανάλυση

- Ποιος είναι ο αριθμός οκτανίων σε ένα μίγμα βενζίνης. Πρόκειται για συγκεκριμένο τύπο βενζίνης?
- Ποια είναι η περιεκτικότητα σε λίπος του νωπού αγελαδινού γάλακτος?
- Ποια είναι η συγκέντρωση Pb στο δίκτυο ύδρευσης μιας πόλης. Είναι πάνω από τα όρια της νομοθεσίας?
- Όταν επαναλαμβάνω μια μέτρηση και παίρνω διαφορετικό αποτέλεσμα, ποιες μετρήσεις κρατάω?
- Ένα δείγμα αναλύεται και δίνει τιμή 2,1 ενώ ένα άλλο τιμή 2,2. Οι τιμές αυτές διαφέρουν μεταξύ τους?
- Ποιος είναι ο χρόνος μιας χημικής αντίδρασης για διαφορετικές θερμοκρασίες. Μπορούμε να καθορίσουμε μια σχέση εξάρτησης?
- Πώς μεταβάλλεται η εκπομπή ρύπων σε σχέση με την παλαιότητα ενός τύπου μηχανής. Μπορούμε να προβλέψουμε τους ρύπους που εκπέμπονται αν γνωρίζουμε το χρόνο λειτουργίας?

Συμπληρωματική ορολογία στη στατιστική

Τυχαία μεταβλητή (τ.μ.) (*random variable*): οποιοδήποτε χαρακτηριστικό του οποίου η τιμή αλλάζει στα διάφορα στοιχεία του πληθυσμού, π.χ. η περιεκτικότητα του γάλακτος σε λίπος

Δεδομένα (*data*): το σύνολο των τιμών που έχουμε στη διάθεσή μας και αφορούν μια τ.μ., π.χ. μετρήσεις της περιεκτικότητας σε λίπος ενός τύπου γάλακτος

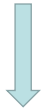
Πληθυσμός (*population*): μια ομάδα ή μια κατηγορία στην οποία αναφέρεται η τ.μ., π.χ. ένας συγκεκριμένος τύπος γάλακτος

(Στατιστικό) Δείγμα (*sample*): ένα υποσύνολο του πληθυσμού που μελετάμε, π.χ. 30 δείγματα γάλακτος του ίδιου τύπου (π.χ. νωπό αγελαδινό)

Παράμετρος (*parameter*): ένα μέγεθος που συνοψίζει με κάποιο τρόπο τις τιμές της τ.μ. στον πληθυσμό, π.χ. η μέση περιεκτικότητα σε λίπος ενός τύπου γάλακτος

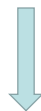
Στατιστικό στοιχείο (*statistic*): ένα μέγεθος που συνοψίζει με κάποιο τρόπο τις τιμές της τ.μ. στο στατιστικό δείγμα, π.χ. η μέση περιεκτικότητα σε λίπος από 30 δείγματα κάποιου τύπου γάλακτος

Γνωρίζουμε τις τιμές της τ.μ.



Συμπεράσματα για την τ.μ.

δείγμα



στατιστικό στοιχείο

πληθυσμός



παράμετρος

Παράμετρος: Η τιμή της αλλάζει και μπορεί να μετρηθεί.

Δύο τύποι παραμέτρων: **συνεχείς (continuous)** και **διακριτές (discrete)**

Συνεχείς: μπορεί να λαμβάνουν άπειρο αριθμό πιθανών τιμών. Περιορίζεται μόνο από τη διαδικασία μέτρησης και την περιοχή στην οποία μπορεί να λαμβάνει τιμές:

Απόσταση 17,4 m

Χρόνος 12,3 s

Μάζα 4,1 Kg

Πιθανότητα 0,05

Συγκέντρωση 12,5 ppm

Διακριτές: μπορεί να λαμβάνουν μόνο συγκεκριμένες τιμές.

Ο αριθμός των δημοσιευμένων εργασιών: 10 (δεν μπορεί να είναι 10,3)

Το στρίψιμο του νομίσματος: 4 (δεν μπορεί να είναι 4,3)

Η απόσταση σε Km με προσέγγιση ακέραιου: 3 (δεν μπορεί να είναι 3,2)

Οι **κατηγορικές** (ή εικονικές παράμετροι) ανήκουν επίσης στις διακριτές παραμέτρους

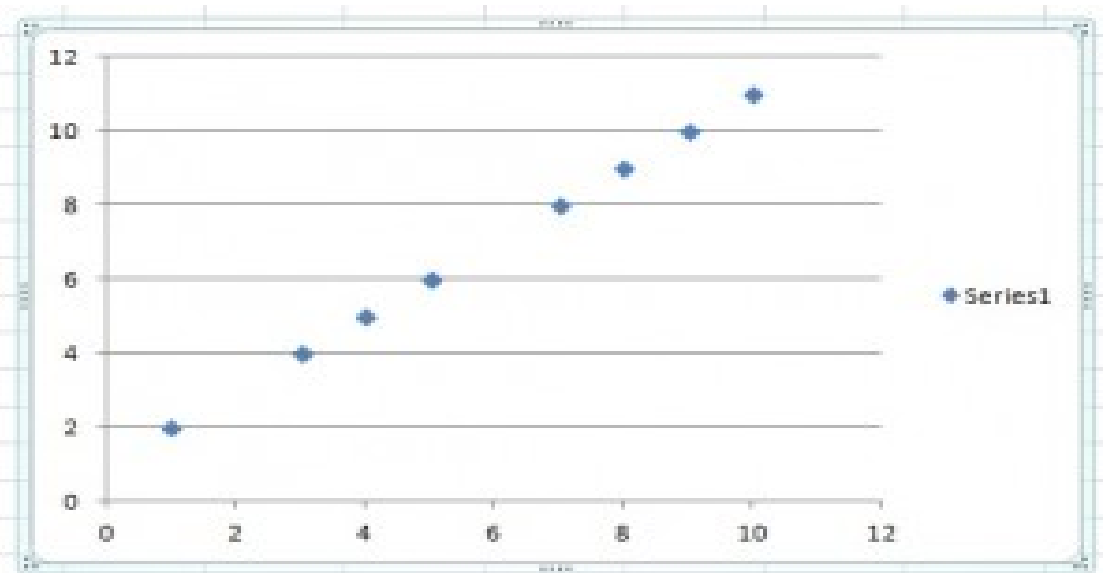
Χρώμα

Οργανοληπτικά χαρακτηριστικά (λίγο γλυκό, γλυκό, πολύ γλυκό)

Σημείωση: Μερικές φορές, αν μια παράμετρος είναι διακριτή ή συνεχής εξαρτάται από τον τρόπο που πρόκειται να χρησιμοποιηθεί στους υπολογισμούς. Η μάζα είναι διακριτή αν αναφερόμαστε σε άτομα ή μόρια. Η ηλικία είναι διακριτή αν και κάποιος μπορεί να είναι 14,3 ετών.

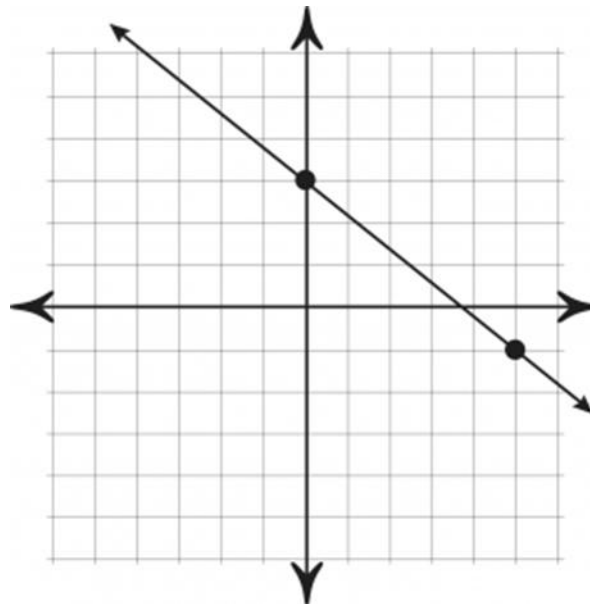
rule of thumb: χρησιμοποιούμε **διακριτές παραμέτρους** μόνο όταν ο αριθμός των επιλογών για τις τιμές τους είναι περιορισμένος με βάση το ερώτημα που απαντάται.

Οι διακριτές παράμετροι είναι όπως τα σημεία σ' ένα σημειόγραμμα (scatterplot) ενώ οι συνεχείς όπως οι γραμμές.



Διακριτά σημεία στο σημειόγραμμα

Συνεχής γραμμή στο γράφημα



it could go on for ever...

ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Περιγραφικά Μέτρα Στατιστικών Δεδομένων

Θα ασχοληθούμε με δύο τύπους περιγραφικών μέτρων:

- τα **μέτρα θέσης-κεντρικής τάσης** (measures of location) που προσδιορίζουν χαρακτηριστικές θέσεις μέσα στο εύρος των δεδομένων και
- τα **μέτρα μεταβλητότητας** (variability measures) που δίνουν περιληπτικά τη διασκόρπιση και μεταβλητότητα των δεδομένων

Μέτρα Θέσης

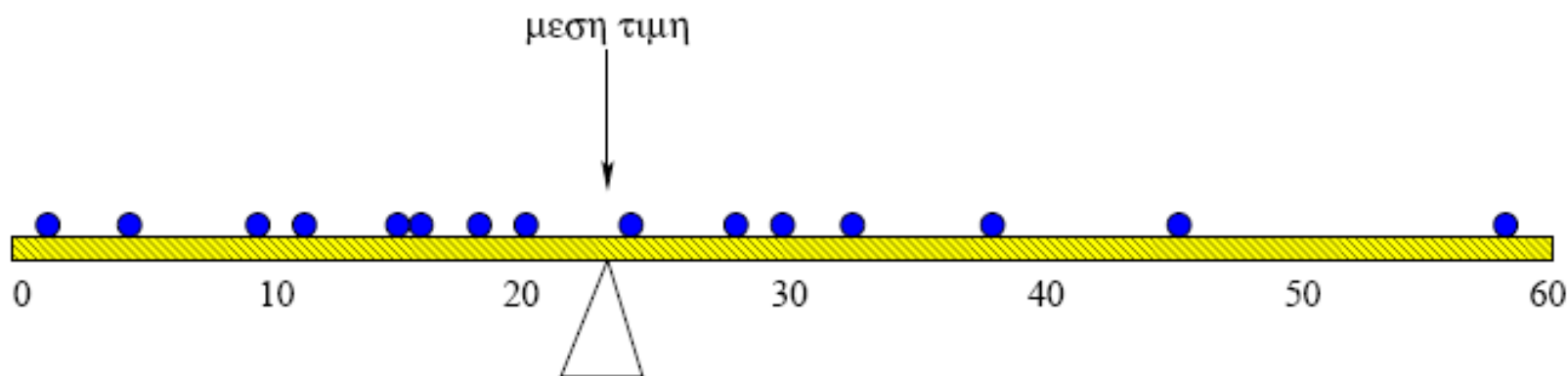
Εννοούμε τα μέτρα **κεντρικής τάσης** που προσδιορίζουν ένα κεντρικό σημείο γύρω από το οποίο τείνουν να συγκεντρώνονται τα δεδομένα. Τα κυριότερα μέτρα θέσης είναι:

- η **δειγματική μέση τιμή** (sample mean value) ή **αριθμητικός μέσος** (arithmetic mean) ή **μέσος όρος** (average)
- η **δειγματική διάμεσος ή διάμεση τιμή** (sample median)
- η **δειγματική επικρατούσα τιμή** (sample mode)

Μέση τιμή: Είναι το πιο γνωστό και χρήσιμο μέτρο του κέντρου των δεδομένων. Έστω $x_1, x_2, \dots, x_n$, οι τιμές των δεδομένων-παρατηρήσεων του δείγματος για μια τυχαία μεταβλητή X που μελετάμε. Η δειγματική μέση τιμή συμβολίζεται $\bar{x}$ κι ορίζεται ως εξής:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Η μέση τιμή είναι το «κέντρο ισορροπίας» των δεδομένων. Για να καταλάβουμε τη φυσική της σημασία, ας φανταστούμε μια σανίδα πάνω στην οποία σκορπίζουμε ένα αριθμό n ίδιων βαριδίων. Το σημείο στήριξης της σανίδας (ώστε να ισορροπεί σε οριζόντια θέση) είναι η μέση τιμή της θέσης των βαριδίων πάνω στη σανίδα.



Σχηματική παρουσίαση της μέσης τιμής.

Διάμεσος: Είναι ένα άλλο μέτρο του κέντρου των δεδομένων και ορίζεται ως η κεντρική τιμή όταν διατάξουμε τα δεδομένα σε αύξουσα σειρά. Θα τη συμβολίσουμε ως $\tilde{x}$.

Αν ο αριθμός n των δεδομένων είναι περιττός τότε η διάμεσος είναι η τιμή στη θέση $(n+1)/2$, ενώ αν ο n είναι άρτιος τότε είναι το ημιάθροισμα των τιμών στις θέσεις $n/2$ και $n/2 + 1$, δηλαδή

$$\tilde{x} = \begin{cases} x_{(n+1)/2} & n = 2k + 1 \\ \frac{x_{n/2} + x_{n/2+1}}{2} & n = 2k \end{cases}$$

Για παράδειγμα, σε δείγμα τριών τιμών η διάμεσος είναι η δεύτερη μικρότερη τιμή και σε δείγμα τεσσάρων είναι ο μέσος όρος της δεύτερης και της τρίτης μικρότερης τιμής.

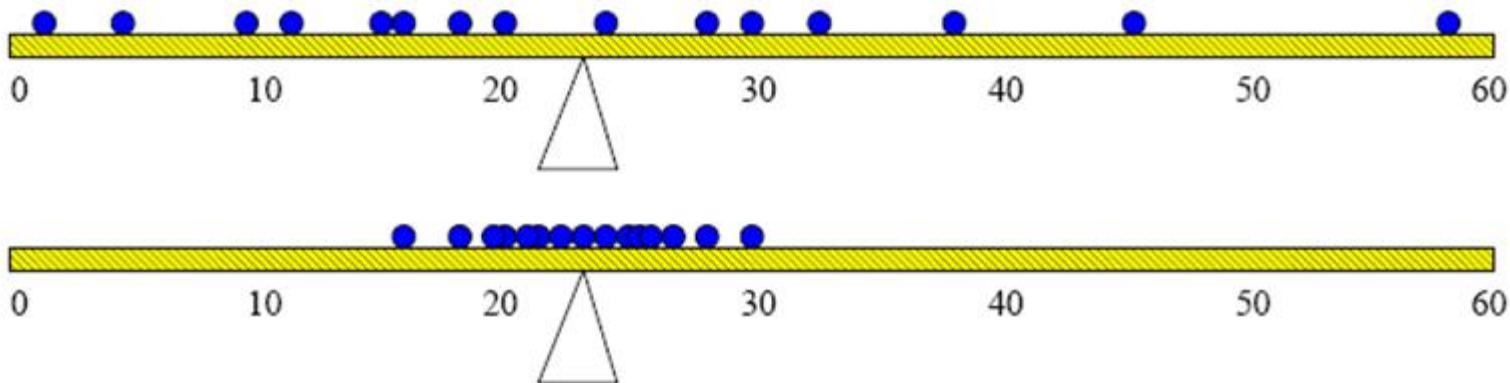
Επικρατούσα τιμή: Χρησιμοποιείται για να δηλώσει την κεντρική τάση των δεδομένων κι ορίζεται ως η τιμή που εμφανίζεται με τη μεγαλύτερη συχνότητα. Αν υπάρχουν πάνω από μια τέτοιες τιμές, τότε όλες αυτές θεωρούνται επικρατούσες τιμές.

Η δειγματική μέση τιμή είναι το πιο σημαντικό από τα τρία μέτρα κεντρικής τάσης. Για τον υπολογισμό της χρησιμοποιούνται όλες οι τιμές του δείγματος ενώ για τη διάμεσο μόνο η τάξη τους. Η μέση τιμή επηρεάζεται από μακρινές τιμές αλλά η διάμεσος όχι (γι αυτό χαρακτηρίζεται και ως **ανθεκτικό στοιχείο-robust statistic**).

Μέτρα μεταβλητότητας

Εκτός από την κεντρική τάση, ενδιαφέρει και η **μεταβλητότητα** ή **διασπορά** των δεδομένων. Όταν τα δεδομένα είναι συγκεντρωμένα γύρω από μια κεντρική τιμή, δηλαδή η διασπορά τους είναι μικρή, τότε η κεντρική τιμή αντιπροσωπεύει ικανοποιητικά τα δεδομένα. Όταν, όμως, τα δεδομένα είναι πολύ διασκορπισμένα, τα μέτρα κεντρικής τιμής δεν δίνουν καλή περιληπτική περιγραφή των δεδομένων.

Δύο δείγματα που έχουν την ίδια μέση τιμή μπορεί να διαφέρουν σημαντικά και κατά χαρακτηριστικό τρόπο ως προς τη διασπορά τους.



Σχηματική παρουσίαση δύο δειγμάτων ίσου πλήθους με ίδια μέση τιμή και διαφορετική μεταβλητότητα.

Κυριότερα **μέτρα μεταβλητότητας** ή **διασποράς**:

- Το **δειγματικό εύρος** (sample range) R
- Η **δειγματική διακύμανση** ή **δειγματική διασπορά** ή **μεταβλητότητα** (sample variance) s^2 και η **δειγματική τυπική απόκλιση** (standard deviation) s
- Τα **εκατοστιαία σημεία** (percentiles) και το **ενδοτεταρτομοριακό εύρος** (interquartile range)

Δειγματικό εύρος (δεδομένων) $R = x_{\max} - x_{\min}$ είναι η διαφορά της ελάχιστης από τη μέγιστη τιμή της τ.μ. του δείγματος.

Υπολογίζεται εύκολα και **δεν είναι ανθεκτικό μέτρο μεταβλητότητας (non-robust statistic)** διότι εξαρτάται μόνο από τις δύο ακραίες παρατηρήσεις $x_{\min}$ και $x_{\max}$ και αγνοεί τις υπόλοιπες.

Μπορεί να αλλάζει σημαντικά από δείγμα σε δείγμα ίδιου πλήθους και από τον ίδιο πληθυσμό.

Γενικά, το εύρος αυξάνει όταν μεγαλώνει το δείγμα καθώς αναμένεται να συμπεριληφθούν πιο ακραίες τιμές.

Διακύμανση ή διασπορά: μετράει τη μεταβλητότητα των δεδομένων-μετρήσεων γύρω από τη μέση τιμή.

Αν ορίσουμε την απόκλιση μιας παρατήρησης x_i και τη μέση τιμή ως $(x_i - \bar{x})$, το άθροισμα όλων αυτών των αποκλίσεων είναι μηδέν διότι χρησιμοποιώντας τον ορισμό της δειγματικής μέσης τιμής έχουμε:

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x} = n\bar{x} - n\bar{x} = 0$$

Η δειγματική μέση τιμή $\bar{x}$ έχει οριστεί έτσι ώστε οι θετικές αποκλίσεις για τιμές μεγαλύτερες του $\bar{x}$ να είναι αθροιστικά ίδιες με τις αρνητικές αποκλίσεις για τιμές μικρότερες του $\bar{x}$.

Για να μετρήσουμε τη μεταβλητότητα των δεδομένων γύρω από τη μέση τιμή επιλέγουμε να αθροίσουμε όχι τις ίδιες τις αποκλίσεις αλλά τα τετράγωνα των αποκλίσεων.

Επίσης, για να ληφθεί ένα μέτρο της μέσης απόκλισης που δεν θα εξαρτάται από το πλήθος των δεδομένων θα πρέπει να διαιρεθεί με το πλήθος τους n .

Για τεχνικούς λόγους (που θα εξηγήσουμε παρακάτω) **διαιρούμε με το $n-1$** αντί για n και η δειγματική διασπορά s^2 ορίζεται ως

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

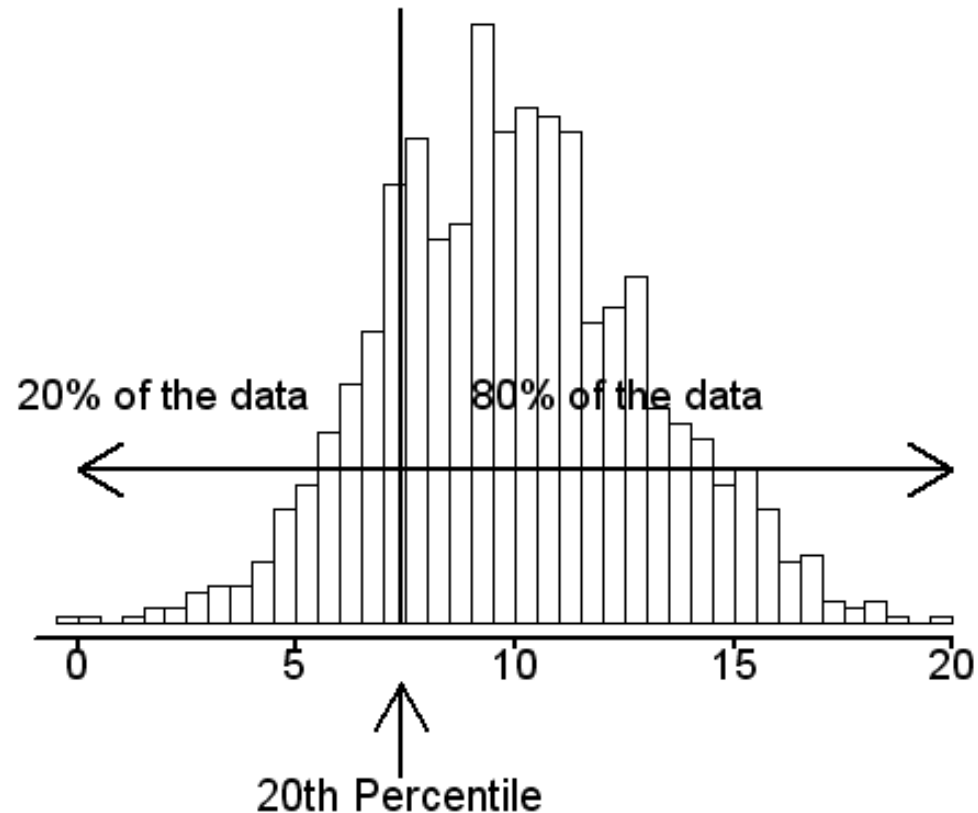
Η **διακύμανση ή μεταβλητότητα s^2** προκύπτει από τα τετράγωνα των μετρήσεων και συχνά είναι δύσκολο να την ερμηνεύσουμε ως πραγματικό φυσικό μέγεθος.

Γι αυτό ορίζουμε τη δειγματική τυπική απόκλιση s που είναι η θετική ρίζα της δειγματικής διασποράς s^2 .

Η **τυπική απόκλιση s** μετριέται με μονάδες μέτρησης της τυχαίας μεταβλητής X κι **εκφράζει** (όπως δηλώνει και η ονομασία της) **την τυπική απόκλιση των δεδομένων από τη δειγματική μέση τιμή, δηλαδή μέχρι πόσο περίπου περιμένουμε μια τυπική τιμή της X να απέχει από τη μέση τιμή.**

Εκατοστιαία σημεία (ή εκατοστημόρια) - ενδοτεταρτομοριακό εύρος. Η διάμεσος χωρίζει τα δεδομένα στα δύο. Μπορούμε να ορίσουμε άλλα σημεία χωρισμού των διατεταγμένων δεδομένων που παίρνουμε από ένα δείγμα. Τέτοια σημεία είναι τα εκατοστιαία σημεία.

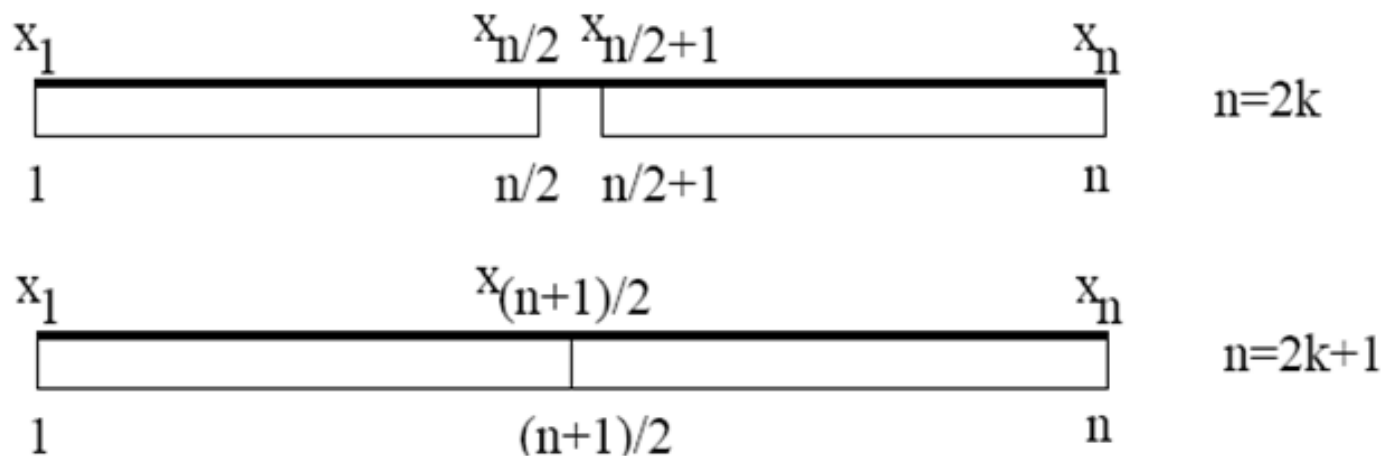
Μια παρατήρηση καλείται p -εκατοστιαίο σημείο (p -percentile) ή p -εκατοστημόριο όταν το ποσοστό μετρήσεων-δεδομένων το πολύ $p\%$ είναι μικρότερο από αυτή την μέτρηση. Η διάμεσος είναι το 50-εκατοστιαίο σημείο.



Άλλα χαρακτηριστικά σημεία είναι αυτά που ορίζουν **τέταρτα** ή **τεταρτημόρια** (quartiles). Το 25-εκατοστιαίο σημείο είναι το **πρώτο** ή **κατώτερο τεταρτημόριο** (first or lower quartile) και συμβολίζεται ως Q_1 , ενώ το 75-εκατοστιαίο σημείο είναι το **τρίτο** ή **ανώτερο τεταρτημόριο** (third or upper quartile) και συμβολίζεται ως Q_3 .

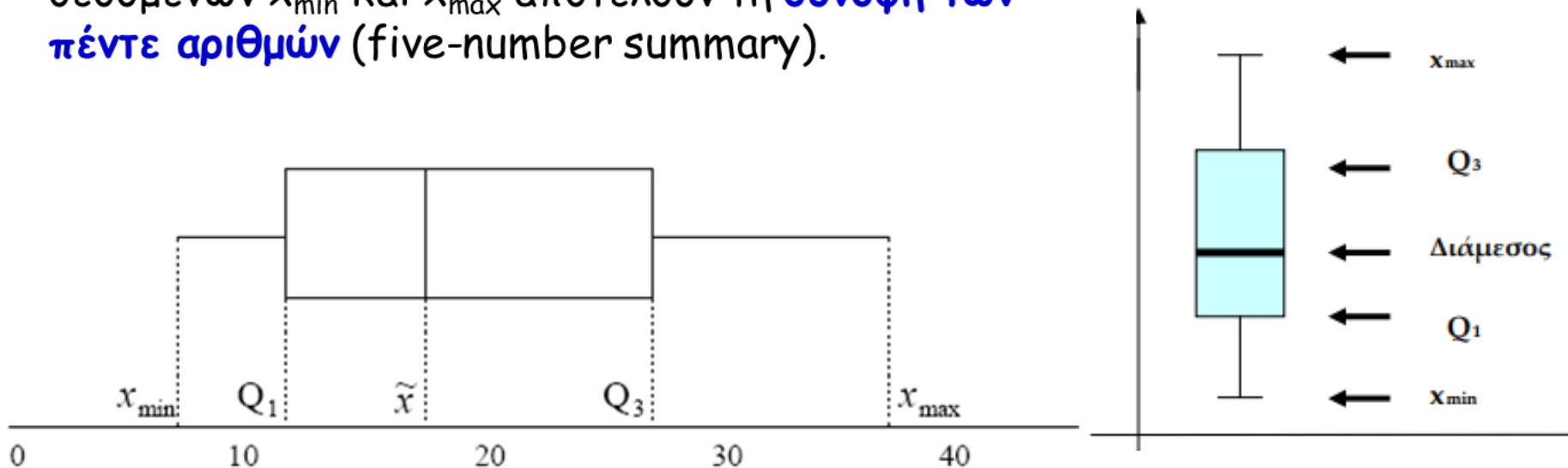
Τα Q_1 και Q_3 ορίζονται όπως η διάμεσος αλλά περιορίζοντας το σύνολο των δεδομένων στα αντίστοιχα υποσύνολα (κατώτερο ή ανώτερο μισό).

Ειδικότερα, έχοντας πρώτα διατάξει τις παρατηρήσεις σε αύξουσα σειρά, αν το σύνολο των n παρατηρήσεων είναι άρτιος αριθμός, τότε το κατώτερο υποσύνολο περιέχει τις παρατηρήσεις 1 έως $n/2$ και το ανώτερο από το $n/2 + 1$ ως n ενώ αν το n είναι περιττός το κατώτερο υποσύνολο περιέχει τις παρατηρήσεις από 1 έως $(n+1)/2$ και το ανώτερο από $(n+1)/2$ ως n .



Η διαφορά $I = Q_3 - Q_1$ λέγεται **ενδοτεταρτομοριακό εύρος** (interquartile range, IQR) και δίνει το εύρος που καλύπτουν τα μισά από τα δεδομένα που είναι πιο κοντά στην κεντρική τιμή (διάμεσο). Το I είναι ένα άλλο μέτρο διασποράς των δεδομένων (**πιο ανθεκτικό-robust** από το εύρος R) που δεν ορίζεται από τη δειγματική μέση τιμή και συνεπώς, δεν επηρεάζεται από αυτήν.

Η διάμεσος $\tilde{x}$, το 1^ο και 3^ο τεταρτημόριο Q_1 και Q_3 αντίστοιχα, καθώς και η ελάχιστη και η μέγιστη τιμή των δεδομένων $x_{\min}$ και $x_{\max}$ αποτελούν τη **σύνοψη των πέντε αριθμών** (five-number summary).



Γραφικά, η παρουσίαση της **σύνοψης των πέντε αριθμών** γίνεται με το **θηκόγραμμα** (box plot) σε οριζόντια ή κάθετη θέση.

Παράδειγμα Θέλουμε να εκτιμήσουμε την περιεκτικότητα σε ραδιενέργεια του χάλυβα που παράγεται από ένα εργοστάσιο Α. Γι αυτό έγιναν μετρήσεις της ραδιενέργειας (σε Bq/g) σε 10 δοκίμια από το εργοστάσιο Α. Τα αποτελέσματα δίνονται στον Πίνακα.

A/A	εργοστάσιο Α	εργοστάσιο Β
1	0.40	0.11
2	0.51	0.13
3	0.51	0.26
4	0.54	0.27
5	0.55	0.33
6	0.59	0.37
7	0.63	0.52
8	0.67	0.65
9	0.75	
10	2.10	
Σύνολο	7.25	2.64

Δεδομένα περιεκτικότητας ραδιενέργειας (σε μονάδα μέτρησης Bq/g) σε δοκίμια από χάλυβα κατασκευασμένα από δύο εργοστάσια Α και Β.

Θα ασχοληθούμε με τις παρατηρήσεις της περιεκτικότητας ραδιενέργειας στο χάλυβα από το εργοστάσιο Α κι ας ονομάσουμε αυτήν την τυχαία μεταβλητή X . Η δειγματική μέση τιμή υπολογίζεται από το άθροισμα των 10 παρατηρήσεων που δίνονται στη δεύτερη στήλη του Πίνακα ως

$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{7.25}{10} = 0.725.$$

Η δειγματική διάμεσος εύκολα μπορεί να βρεθεί μια κι οι παρατηρήσεις δίνονται σε αύξουσα σειρά. Αφού το n είναι άρτιο η διάμεσος δίνεται ως

$$\tilde{x} = \frac{x_{n/2} + x_{n/2+1}}{2} = \frac{x_5 + x_6}{2} = \frac{0.55 + 0.59}{2} = 0.57.$$

Παρατηρούμε ότι η δειγματική μέση τιμή δίνει μεγαλύτερη τιμή στην εκτίμηση της κεντρικής τάσης των δεδομένων από τη δειγματική διάμεσο. Αυτό συμβαίνει γιατί η μέση τιμή επηρεάζεται από την ακραία τιμή της δεκάτης παρατήρησης που είναι πολύ μεγαλύτερη από όλες τις άλλες.

Θα πρέπει να γνωρίζουμε αν αυτή η ακραία τιμή είναι πραγματική ή οφείλεται σε κάποιο σφάλμα της παρατήρησης

Η μικρότερη περιεκτικότητα ραδιενέργειας στο δείγμα είναι $x_{\min} = 0.40 \text{ Bq/g}$ και η μεγαλύτερη είναι $x_{\max} = 2.10 \text{ Bq/g}$. Άρα εύρος δεδομένων $R = 1.70 \text{ Bq/g}$,
είναι μεγάλο εξαιτίας της ακραίας τιμής που συμπεριλάβαμε στο δείγμα.

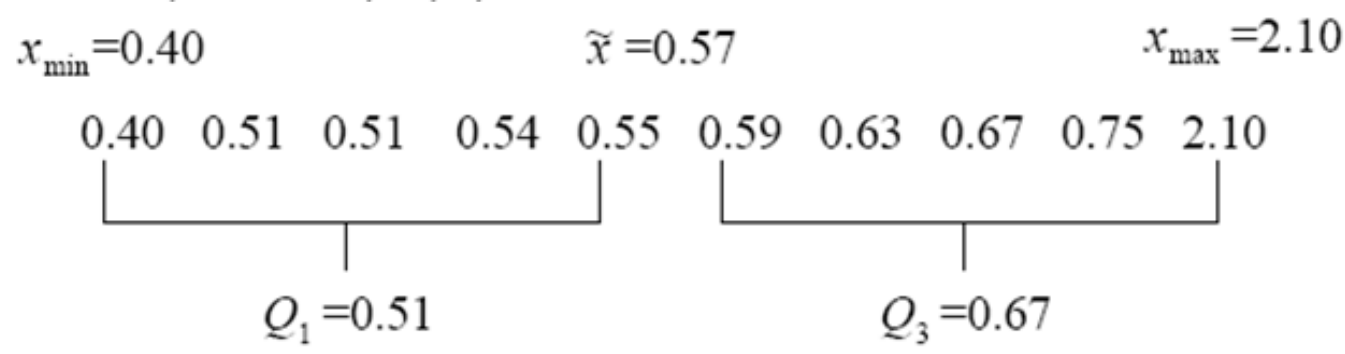
Για να βρούμε τη διασπορά s^2 της περιεκτικότητας της ραδιενέργειας στο χάλυβα στο δείγμα αντικαθιστώντας το στον τύπο της δειγματικής διασποράς βρίσκουμε $s^2 = 0.243$.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Η δειγματική τυπική απόκλιση είναι $s = \sqrt{s^2} = \sqrt{0.243} = 0.493$,

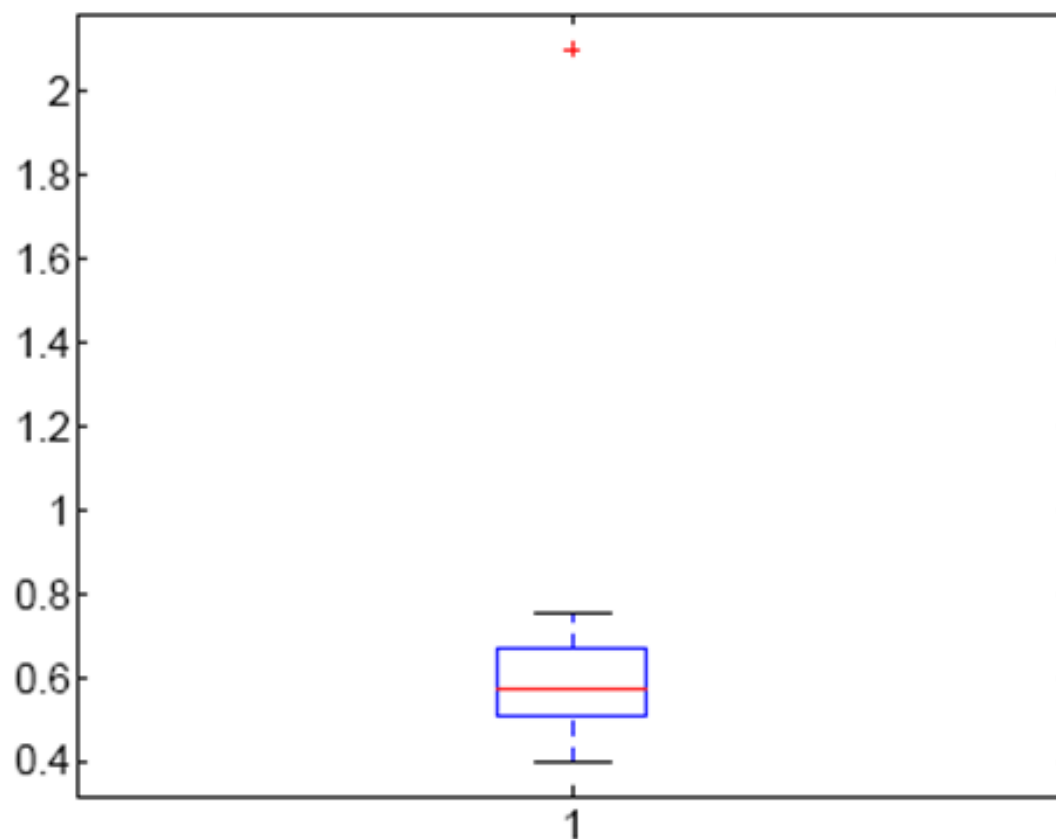
δηλαδή η αντιπροσωπευτική τυπική απόκλιση από τη μέση περιεκτικότητα ραδιενέργειας είναι περίπου 0.5 Bq/g , που είναι πολύ μεγάλη (όση περίπου και η διάμεσος).

Η σύνοψη των 5 αριθμών δίνεται στο Σχήμα όπου παρουσιάζεται σχηματικά και η εύρεση του πρώτου και τρίτου τεταρτομορίου.



Από το πρώτο και τρίτο τεταρτομόριο υπολογίζουμε $I = Q_3 - Q_1 = 0.67 - 0.51 = 0.16 \text{ Bq/g}$. Έχοντας βρεί τη σύνοψη των 5 αριθμών μπορούμε εύκολα να παραστήσουμε το θηκόγραμμα.

θηκόγραμμα σε κατακόρυφη θέση. Παρατηρούμε ότι ο άνω μύστακας δεν προεκτείνεται ως τη μέγιστη τιμή των δεδομένων που είναι 2.10. Αυτή η τιμή δηλώνεται ως ακραία τιμή με ιδιαίτερο σύμβολο, αφού η απόσταση του αντίστοιχου σημείου από το πάνω μέρος του κουτιού, $Q_3 = 0.67$, είναι μεγαλύτερη από $3I = 3 \cdot 0.16 = 0.48$.



Θηκόγραμμα των 10 παρατηρήσεων περιεκτικότητας ραδιενέργειας σε χάλυβα του εργοστασίου Α.

Παράδειγμα Θεωρούμε πως η ακραία τιμή στο προηγούμενο παράδειγμα οφείλεται πράγματι σε κάποιο σφάλμα της μέτρησης και την απαλείφουμε. Για το νέο δείγμα των 9 παρατηρήσεων υπολογίζουμε τα μέτρα θέσης και μεταβλητότητας.

Η μέση τιμή είναι

$$\bar{x} = \frac{1}{9} \sum_{i=1}^9 x_i = \frac{5.15}{9} = 0.572$$

και η διάμεσος (το n είναι περιττό)

$$\tilde{x} = x_{(n+1)/2} = x_5 = 0.55.$$

Βλέπουμε λοιπόν ότι μετά την απαλοιφή της ακραίας τιμής η δειγματική μέση τιμή και η διάμεσος δίνουν περίπου την ίδια εκτίμηση της κεντρικής τάσης των δεδομένων, που είναι ένδειξη συμμετρίας της κατανομής.

Για τα άλλα μέτρα έχουμε:

$$\text{Διασπορά} \quad s^2 = \frac{1}{8} (5.15 - 9 \cdot 0.572^2) = 0.010.$$

$$\text{Τυπική απόκλιση} \quad s = \sqrt{0.010} = 0.10$$

$$\text{Ελάχιστη τιμή} \quad x_{\min} = 0.40$$

$$\text{Μέγιστη τιμή} \quad x_{\max} = 0.75$$

$$\text{Εύρος} \quad R = 0.75 - 0.40 = 0.35$$

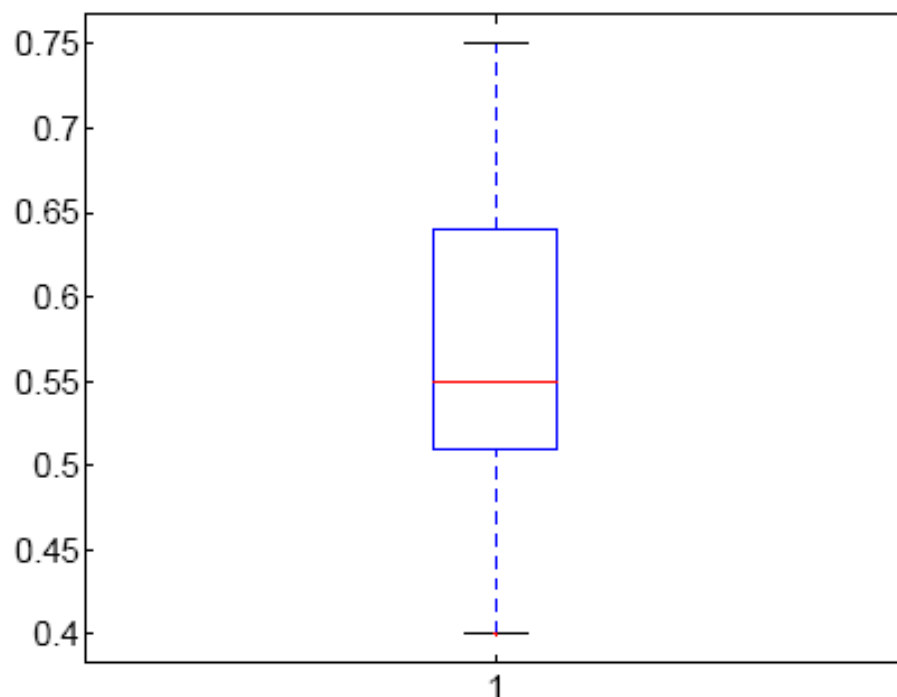
$$\text{Πρώτο τεταρτομόριο} \quad (\text{διάμεσος των } \{x_1, \dots, x_5\}) \quad Q_1 = x_3 = 0.51$$

$$\text{Τρίτο τεταρτομόριο} \quad (\text{διάμεσος των } \{x_5, \dots, x_9\}) \quad Q_3 = x_7 = 0.63$$

$$\text{Ενδοτεταρτομοριακό εύρος} \quad I = 0.63 - 0.51 = 0.12$$

Το αντίστοιχο θηκόγραμμα δίνεται στο Σχήμα .

Θηκόγραμμα περιεκτικότητας ραδιενέργειας



Θηκόγραμμα των 9 παρατηρήσεων περιεκτικότητας ραδιενέργειας σε χάλυβα του εργοστασίου Α.

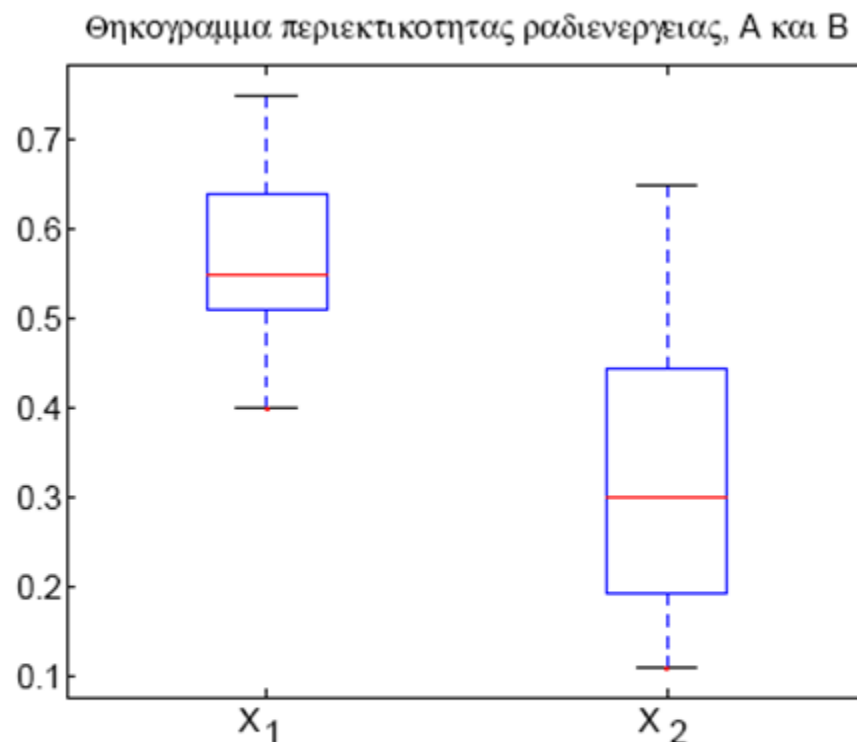
Από τα μέτρα θέσης και μεταβλητότητας καθώς κι από το θηκόγραμμα παρατηρούμε ότι με την απαλοιφή της ακραίας τιμής η κατανομή της περιεκτικότητας ραδιενέργειας στο χάλυβα του εργοστασίου Α φαίνεται να είναι συμμετρική και μπορούμε να κάνουμε τώρα την παραδοχή ότι η κατανομή είναι κανονική με βάση το δείγμα.

Παράδειγμα Στον Πίνακα παρουσιάζονται επίσης οι μετρήσεις της περιεκτικότητας ραδιενέργειας σε 8 δοκίμια χάλυβα που κατασκευάστηκαν από ένα άλλο εργοστάσιο B. Θέλουμε να συγκρίνουμε την περιεκτικότητα ραδιενέργειας στο χάλυβα από τις δύο μονάδες παραγωγής με βάση τα δείγματα των 9 και 8 παρατηρήσεων που έχουμε από το εργοστάσιο A και από το εργοστάσιο B αντίστοιχα. Έστω X_1 η τ.μ. της περιεκτικότητας ραδιενέργειας στο χάλυβα για το εργοστάσιο A και X_2 η αντίστοιχη τ.μ. για το εργοστάσιο B. Στο προηγούμενο παράδειγμα υπολογίσαμε τα μέτρα θέσης και μεταβλητότητας για τη X_1 . Κάνουμε το ίδιο για τη X_2 . Τα συγκεντρωτικά αποτελέσματα δίνονται στον Πίνακα

Πίνακας. Μέτρα θέσης και μεταβλητότητας για τα δεδομένα περιεκτικότητας ραδιενέργειας σε 9 και 8 δοκίμια χάλυβα κατασκευασμένα από τα εργοστάσια A και B στη στήλη 2 και 3 αντίστοιχα.

Μέτρο	X_1	X_2
Μέση τιμή	$\bar{x}_1 = 0.572$	$\bar{x}_2 = 0.33$
Διάμεσος	$\tilde{x}_1 = 0.55$	$\tilde{x}_2 = 0.30$
Διασπορά	$s_1^2 = 0.010$	$s_2^2 = 0.034$
Τυπική απόκλιση	$s_1 = 0.10$	$s_2 = 0.18$
Ελάχιστη τιμή	$x_{1,\min} = 0.40$	$x_{2,\min} = 0.11$
Μέγιστη τιμή	$x_{1,\max} = 0.75$	$x_{1,\max} = 0.65$
Εύρος	$R_1 = 0.35$	$R_2 = 0.54$
Πρώτο τεταρτομόριο	$Q_{1,1} = 0.51$	$Q_{2,1} = 0.195$
Τρίτο τεταρτομόριο	$Q_{1,3} = 0.63$	$Q_{2,3} = 0.445$
Ενδοτεταρτομοριακό εύρος	$I_1 = 0.12$	$I_2 = 0.250$

Επίσης στο Σχήμα δίνεται το συνδυασμένο θηκόγραμμα για το δείγμα περιεκτικότητας ραδιενέργειας σε χάλυβα κι από τα δύο εργοστάσια.



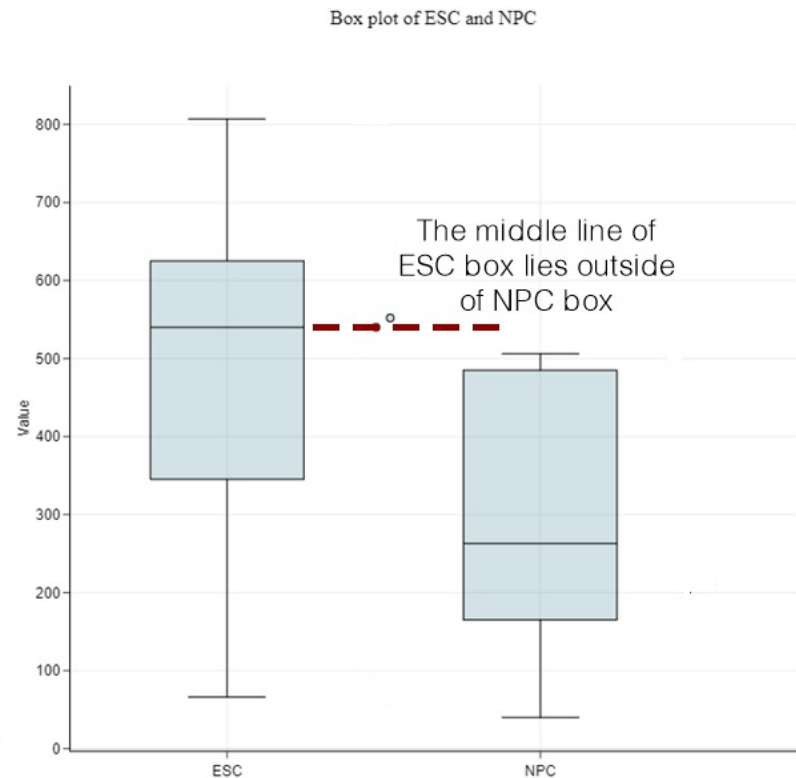
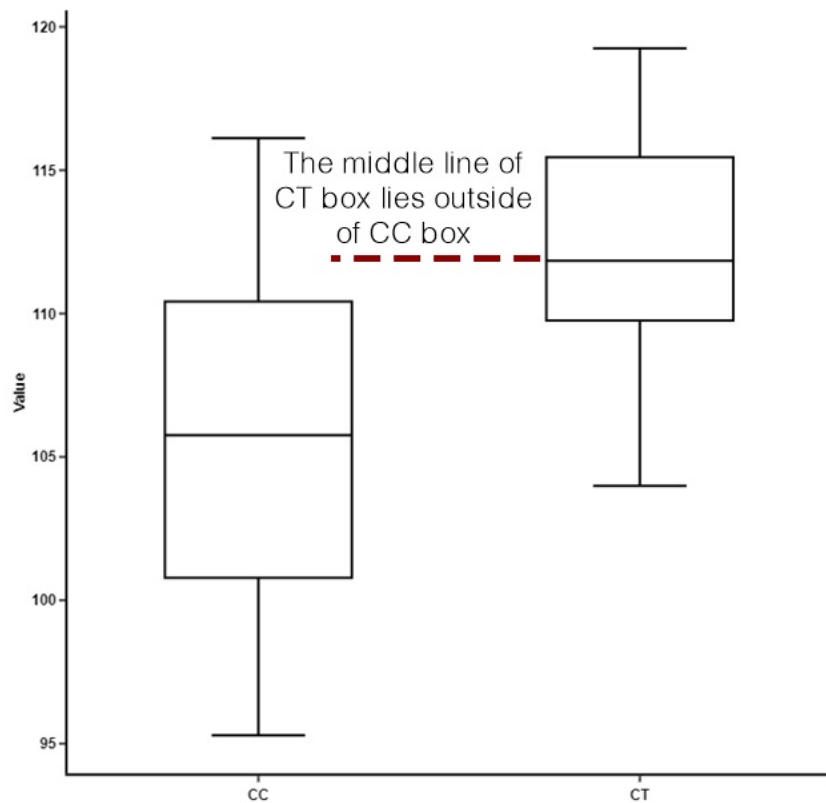
Θηκόγραμμα των 9 παρατηρήσεων περιεκτικότητας ραδιενέργειας σε χάλυβα του εργοστασίου A και των 8 παρατηρήσεων περιεκτικότητας ραδιενέργειας σε χάλυβα του εργοστασίου B.

Από τα μέτρα θέσης (μέση τιμή και διάμεσο) φαίνεται ότι η κεντρική τάση της περιεκτικότητας ραδιενέργειας είναι μεγαλύτερη για το χάλυβα του εργοστασίου A. Από τα μέτρα μεταβλητότητας (τυπική απόκλιση, εύρος δεδομένων και ενδοτεταρτομοριακό εύρος) φαίνεται πως η περιεκτικότητα ραδιενέργειας μεταβάλλεται λιγότερο στο χάλυβα του εργοστασίου A.

How to compare box plots

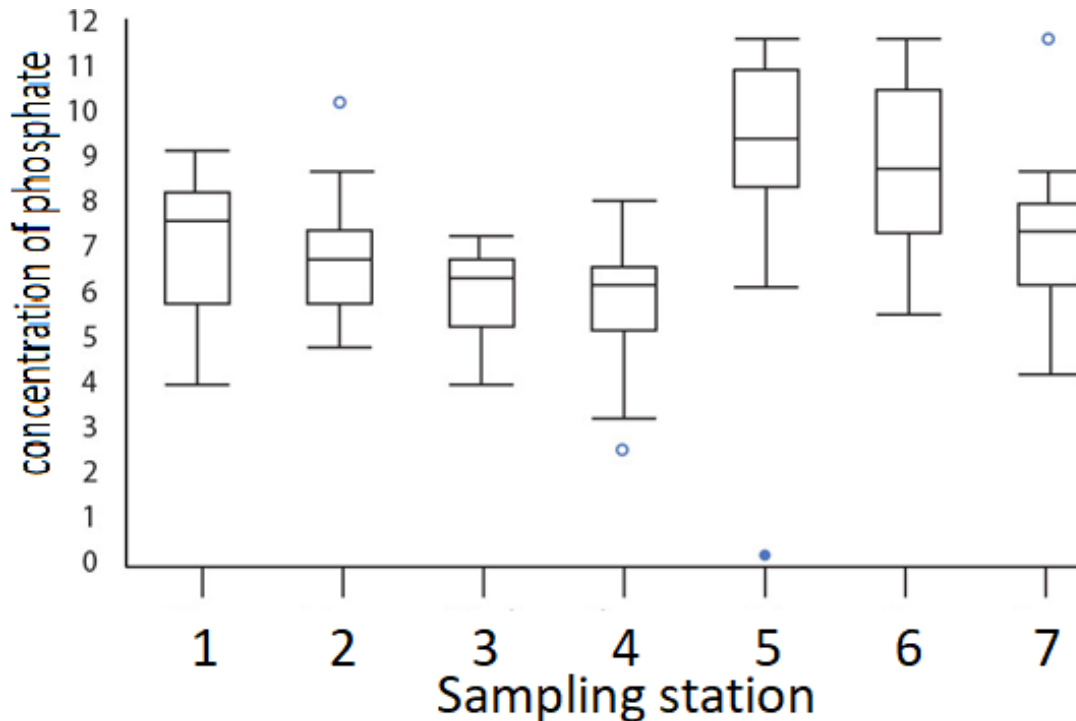
Step 1: Compare the medians of box plots

Compare the respective medians of each box plot. If the median line of a box plot lies outside the box of a comparison box plot, then there is likely a difference between the two groups.



Step 2: Compare the interquartile ranges and whiskers of box plots

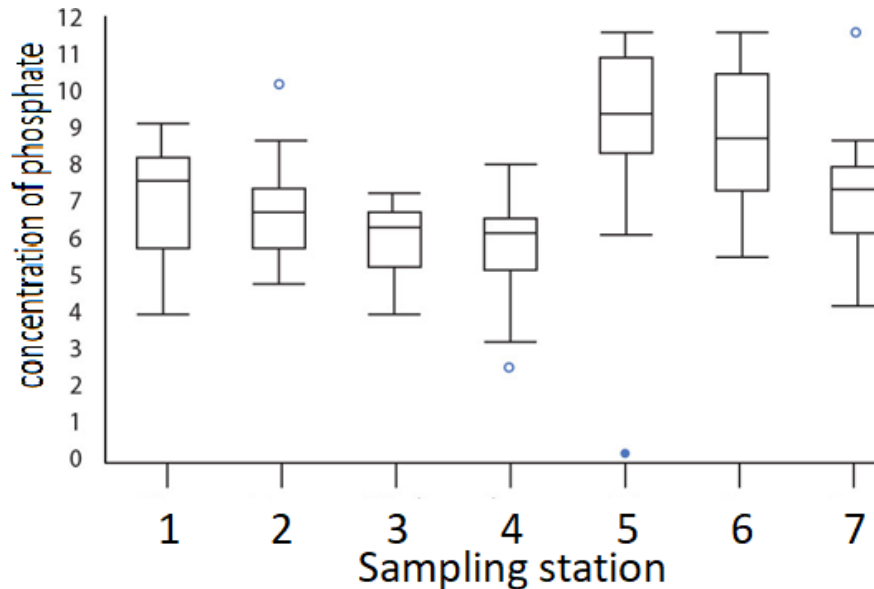
Compare the interquartile ranges (that is, the box lengths), to examine how the data are dispersed between each sample. The longer the box the more dispersed the data. The smaller the box the less dispersed the data.



Next, look at the overall spread as shown by the values at the end of two whiskers.

Step 3: Look for potential outliers

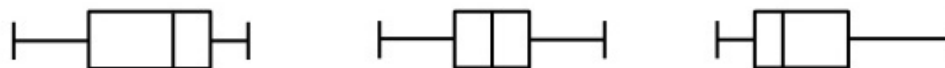
When reviewing a box plot, an outlier is defined as a data point that is located outside the whiskers of the box plot.



Step 4: Look for signs of skewness

If the data do not appear to be symmetric, does each sample show the same kind of asymmetry?

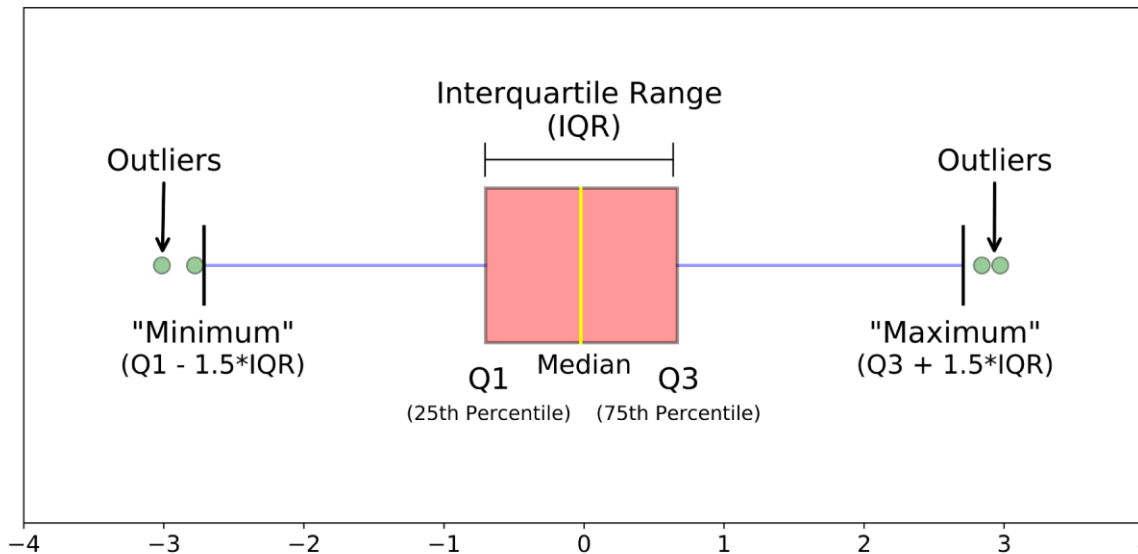
ΘΑ ΑΝΑΦΕΡΘΟΥΜΕ ΣΤΗ ΣΥΝΕΧΕΙΑ



Box plots are useful as they show outliers (ακραίες τιμές) within a data set.

An outlier is an observation that is numerically distant from the rest of the data.

When reviewing a box plot, an outlier is defined as a data point that is located outside the whiskers of the box plot.



For example, outside **1.5 times** the interquartile range above the upper quartile and below the lower quartile ($Q_1 - 1.5 * IQR$ or $Q_3 + 1.5 * IQR$).

Αυτό είναι ένα άλλο κριτήριο εύρεσης ακραίων τιμών

Η έννοια του **πληθυσμού** στη στατιστική συνδέεται με το σύνολο των τιμών μιας τυχαίας μεταβλητής X .

Αν αυτό δεν είναι εφικτό (σχεδόν αδύνατο!) γιατί οι πληθυσμοί είναι μεγάλοι, πραγματοποιούμε μετρήσεις που αποτελούν ένα (στατιστικό) **δείγμα** από τον παραπάνω πληθυσμό. Έχει σημασία το δείγμα:

1. να προέρχονται από τη μελέτη ενός πληθυσμού
2. να έχει ληφθεί τυχαία (τα πειραματικά δεδομένα που αποτελούν το δείγμα να είναι αντιπροσωπευτικά του πληθυσμού όλων των πειραματικών δεδομένων) και
3. να οδηγεί σε ασφαλή συμπεράσματα για τις ιδιότητες (της μεταβλητής X) του πληθυσμού από τον οποίο προέρχεται (το δείγμα).

Ένας συνηθισμένος τρόπος παρουσίασης των αποτελεσμάτων μιας σειράς μετρήσεων μιας μεταβλητής X ενός πληθυσμού είναι οι **κατανομές συχνοτήτων**.

Συχνότητα είναι ο αριθμός των τιμών ενός πληθυσμού που ανήκουν σε μία **τάξη**.

Οι σπουδαιότερες κατανομές που αφορούν τη μελέτη τυχαίων σφαλμάτων κατά τις μετρήσεις φυσικών μεγεθών είναι η διωνυμική κατανομή, η κατανομή Poisson και η κατανομή Gauss ή κανονική κατανομή.

ΠΕΡΙΓΡΑΦΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Οι στατιστικοί πίνακες και γραφικές παραστάσεις αποτελούν χρήσιμα εργαλεία για την παρουσίαση των δεδομένων καθαρά, σύντομα και με σαφήνεια. Μπορούν να αποκαλύψουν σημαντικά χαρακτηριστικά, όπως το εύρος, τη συμμετρικότητα των δεδομένων και την ύπαρξη ακραίων τιμών.

Πίνακας συχνοτήτων. Τα δεδομένα ενός δείγματος για μία τ.μ. X που παίρνει τιμές σε ένα σχετικά μικρό σύνολο τιμών μπορούν να παρουσιαστούν σε ένα **πίνακα συχνοτήτων** (frequency table). Ο πίνακας παρουσιάζει για κάθε τιμή x_i της X τη συχνότητα εμφάνισής της, f_i , δηλαδή, πόσες φορές εμφανίζεται η κάθε τιμή στο δείγμα.

Εύκολα μπορούμε να υπολογίσουμε και τη **σχετική συχνότητα** (relative frequency) εμφάνισης ή αλλιώς το ποσοστό (percent) p_i που ορίζεται από το λόγο της συχνότητας εμφάνισης f_i μίας τιμής x_i προς το σύνολο των μετρήσεων n του δείγματος

$$p_i = \frac{f_i}{n}$$

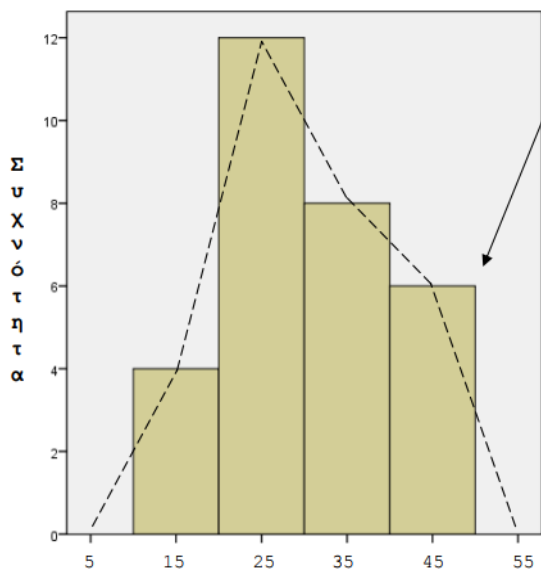
Ορίζεται και η **αθροιστική συχνότητα** (cumulative frequency) F_i μίας τιμής x_i ως το άθροισμα των συχνοτήτων όλων των τιμών που είναι μικρότερες ή ίσες του x_i ,

$$F_i = \sum_{j=1}^i f_j \quad \text{όπου } x_j \leq x_i \quad \text{για } j \leq i$$

Με τον ίδιο τρόπο ορίζεται και η σχετική αθροιστική συχνότητα P_i μίας τιμής x_i ως το άθροισμα των συχνοτήτων όλων των τιμών που είναι μικρότερες ή ίσες του x_i ,

$$P_i = \sum_{j=1}^i p_j \quad \text{όπου } x_j \leq x_i \quad \text{για } j \leq i$$

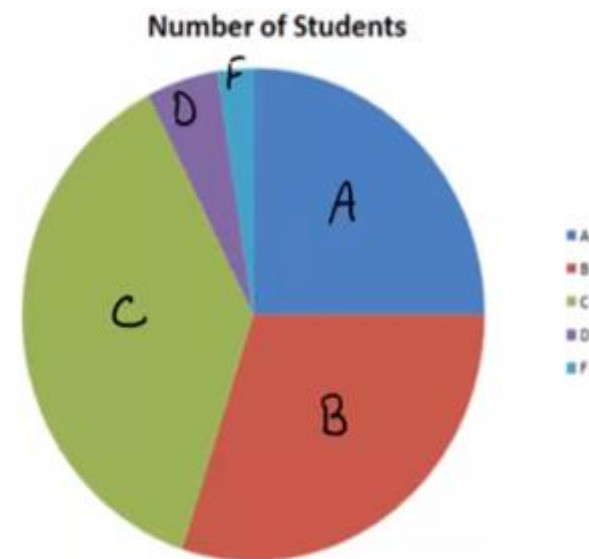
Ραβδόγραμμα. Τα δεδομένα ενός πίνακα συχνοτήτων εύκολα μπορούν να παρασταθούν γραφικά σ ένα ραβδόγραμμα (bar chart), όπου η κάθε ράβδος παρουσιάζει τη συχνότητα για κάθε τιμή x_i . Η ίδια πληροφορία μπορεί να δοθεί και με άλλου είδους γραφήματα, όπως μ ένα **κυκλικό διάγραμμα** ή **διάγραμμα πίτας** (pie chart) όπου το κάθε κομμάτι της επιφάνειας του κύκλου (πίτα) παρουσιάζει τη συχνότητα της αντίστοιχης τιμής.



Πολύγωνο συχνοτήτων

Κατασκευάζεται ενώνοντας με ευθείες γραμμές τα μέσα των πάνω πλευρών των ορθογώνιων του ιστογράμματος.

Grade	Frequency
A	10
B	12
C	15
D	2
F	1
40	



Παράδειγμα Μια εταιρεία τροφοδοσίας χημικών προϊόντων πουλάει ένα χημικό προϊόν στους πελάτες της σε παρτίδες των 5 λίτρων. Η εταιρεία θέλει να μελετήσει την ποσότητα του προϊόντος σε κάθε παραγγελία. Από τα αρχεία της εταιρείας βρέθηκαν 120 πρόσφατες παραγγελίες αυτού του χημικού προϊόντος και ο αριθμός των παρτίδων σε κάθε παραγγελία δίνεται στον Πίνακα

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

Η τ.μ. X είναι ο αριθμός των παρτίδων του χημικού προϊόντος ανά παραγγελία. Τα δεδομένα του Πίνακα είναι διακριτά αριθμητικά (ως αριθμοί από το 1 ως το ανώτατο αριθμό παρτίδων). Με κατάλληλη επεξεργασία θα μπορούσαμε να οργανώσουμε τα δεδομένα σε διατακτικά κατηγορικά, δηλαδή ως κατηγορίες μεγέθους παραγγελίας με βάση τον αριθμό των παρτίδων, μικρού μεγέθους για 1 ή 2 παρτίδες, μεσαίου μεγέθους για 3 ή 4 παρτίδες και μεγάλου μεγέθους για περισσότερες από 4 παρτίδες.

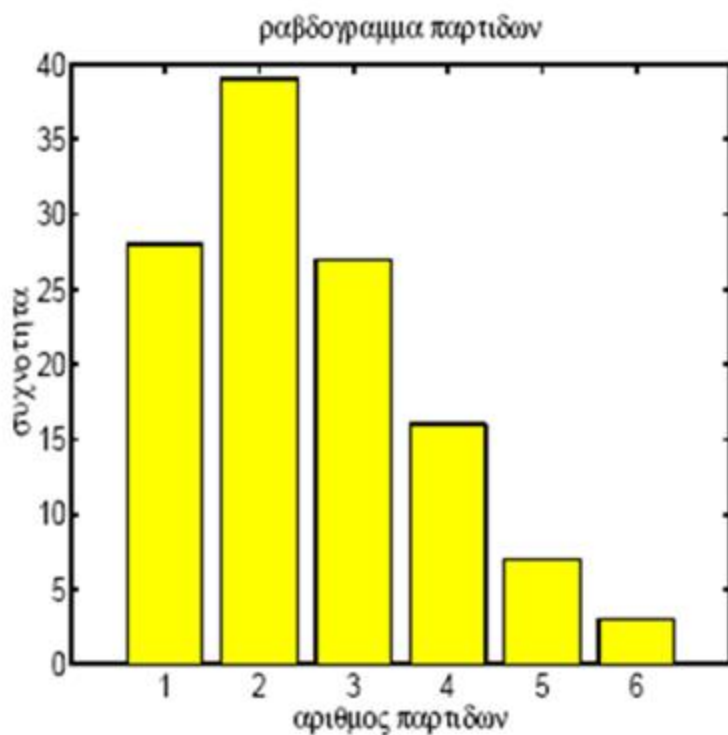
Με βάση τον Πίνακα δεν είναι εύκολο να μελετήσουμε την συχνότητα εμφάνισης των δι-αφόρων αριθμών παρτίδων. Ποιος αριθμός παρτίδων εμφανίζεται συχνότερα; Είναι περισσότερες παραγγελίες των 2 παρτίδων ή των 3 παρτίδων; Για να απαντήσουμε σε τέτοια ερωτήματα μπορούμε να μετρήσουμε πόσες φορές εμφανίζεται ο κάθε αριθμός παρτίδων και να φτιάξουμε έτσι τον πίνακα συχνοτήτων. Αυτούς τους απλούς υπολογισμούς μπορούμε εύκολα να τους κάνουμε μόνοι μας αλλά όταν το δείγμα είναι μεγάλο θα χρειαστεί να χρησιμοποιήσουμε κάποιο υπολογιστικό πρόγραμμα

Πίνακας συχνοτήτων για τον αριθμό παρτίδων σε 120 παραγγελίες που περιλαμβάνει τη συχνότητα f_i , τη σχετική συχνότητα p_i , τη θροιστική συχνότητα F_i και τη σχετική θροιστική συχνότητα P_i).

x_i	f_i	p_i	F_i	P_i
1	28	0.23	28	0.23
2	39	0.33	67	0.56
3	27	0.23	94	0.78
4	16	0.13	110	0.92
5	7	0.06	117	0.97
6	3	0.03	120	1.00
Άθροισμα	120	1.00		

Ο Πίνακας παρουσιάζει τις τιμές της X (αριθμός παρτίδων x_i για $i = 1, \dots, 6$) στην πρώτη στήλη, τη συχνότητα f_i της κάθε τιμής x_i στη δεύτερη στήλη, τη σχετική συχνότητα (ποσοστό) p_i στην τρίτη στήλη, την θροιστική συχνότητα F_i στην τέταρτη στήλη και τη σχετική θροιστική συχνότητα P_i στην πέμπτη στήλη.

Στο Σχήμα παρουσιάζεται το ραβδόγραμμα που προκύπτει από τις συχνότητες εμφάνισης του κάθε αριθμού παρτίδων. Από τον πίνακα συχνοτήτων και το ραβδόγραμμα είναι φανερό πως οι



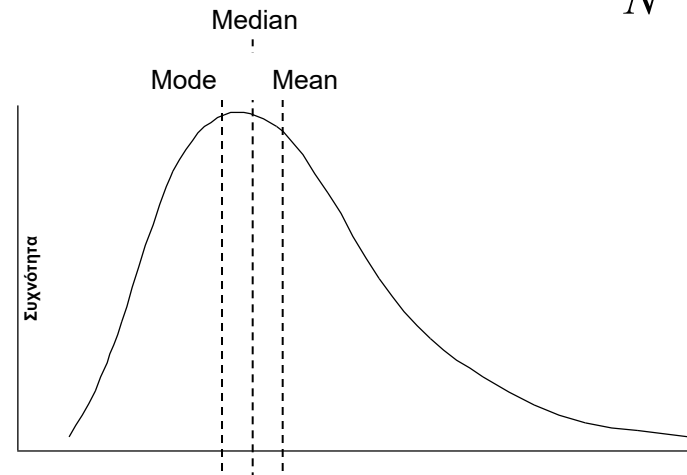
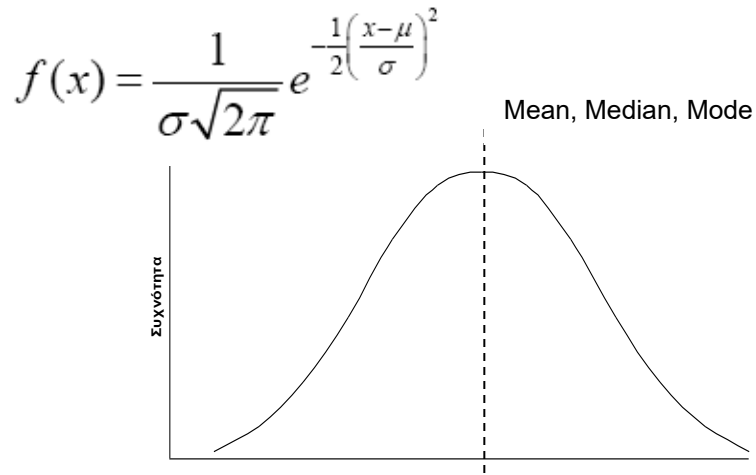
Ραβδόγραμμα του αριθμού παρτίδων σε 120 παραγγελίες του Πίνακα

περισσότερες παραγγελίες στο δείγμα μας είναι 2 παρτίδων, λιγότερες είναι 1 παρτίδας ή 3 παρτίδων και η συχνότητα φθίνει καθώς αυξάνει ο αριθμός παρτίδων, δηλαδή για παραγγελίες 4, 5 και 6 παρτίδων.

Κατανομή Gauss ή κανονική κατανομή (στην ανάλυση)

Αν μία μέτρηση επαναλαμβάνεται κάτω από ίδιες συνθήκες, τα αποτελέσματα κάθε μέτρησης, x , θα κατανέμονται τυχαία γύρω από μία μέση τιμή εξαιτίας ανεξέλεγκτων ή πειραματικών σφαλμάτων. Αν μπορεί να ληφθούν άπειρες τέτοιες μετρήσεις, αυτές θα κατανέμονται σε μια καμπύλη όμοια μ' εκείνες του Σχήματος. Η αριστερή καμπύλη απεικονίζει μια κανονική ή *Gaussian* κατανομή που περιγράφεται με ακρίβεια από τη μέση τιμή του πληθυσμού μ (IUPAC) και την τυπική απόκλιση σ .

Η μέση τιμή του πληθυσμού μ είναι το άθροισμα όλων των τιμών N ως:
$$\mu = \frac{\sum x_i}{N}$$



Τυπική απόκλιση σ μιας κανονικής κατανομής:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{N}}{N}}$$

Επειδή, όμως, άπειρος αριθμός μετρήσεων ποτέ δεν είναι δυνατόν να ληφθεί, μια εκτίμηση του μ γίνεται χρησιμοποιώντας την ίδια διαδικασία άθροισης με n ισοδύναμο με πεπερασμένο αριθμό μετρήσεων (π.χ. 10, 20 κλπ). Η εκτίμηση αυτή του μ ορίζεται ως $\bar{x}$.

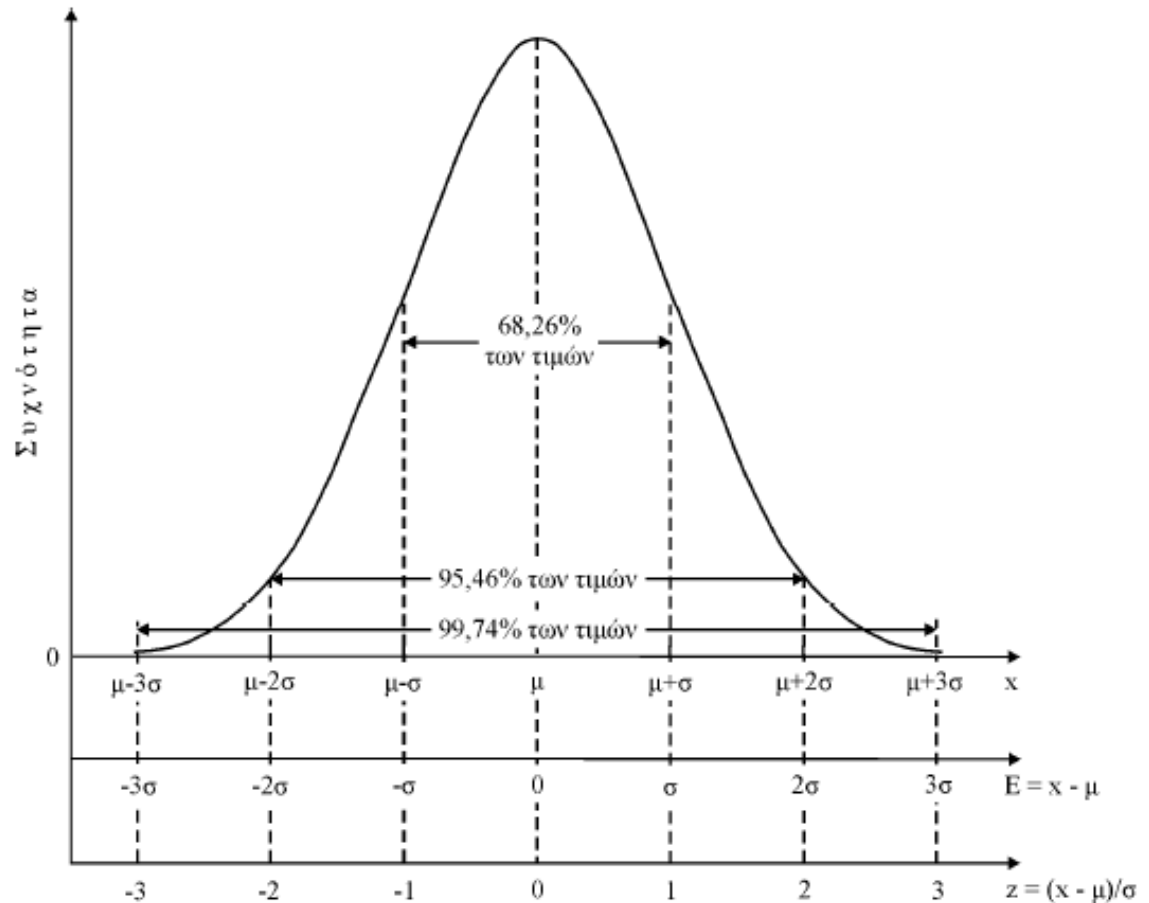
Ο αναλυτής μπορεί να εκτιμήσει την τυπική απόκλιση από τις άπειρες μετρήσεις που ελήφθησαν.

Μια εκτίμηση του σ δίνεται, ωστόσο, από το μέγεθος s για n μετρήσεις που υπολογίζεται ως

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}}$$

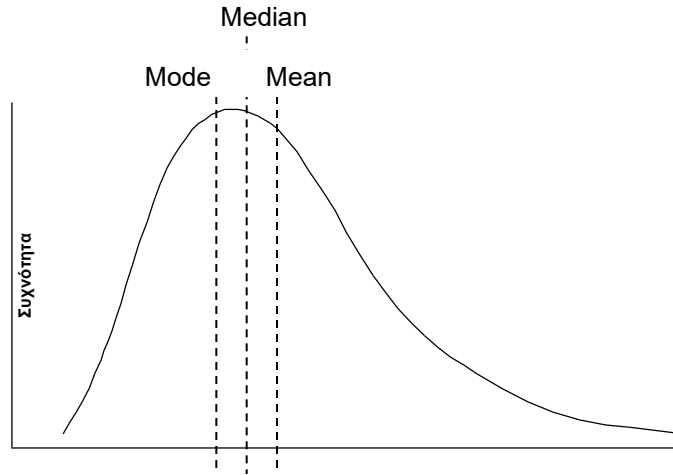
Η τυπική απόκλιση καθορίζει τη διασπορά της κανονικής κατανομής και περιλαμβάνει ένα συγκεκριμένο κλάσμα τιμών που δημιουργούν την καμπύλη.

Το 68,26 % των τιμών βρίσκονται μεταξύ του $\mu \pm 1\sigma$, το 95,46 % μεταξύ $\mu \pm 2\sigma$ και το 99,74 % μεταξύ $\mu \pm 3\sigma$. Είναι ακριβές να θεωρηθεί ότι το 95 % των τιμών είναι στο διάστημα $\pm 2\sigma$ και το 99 % των τιμών στο διάστημα $\pm 3\sigma$.



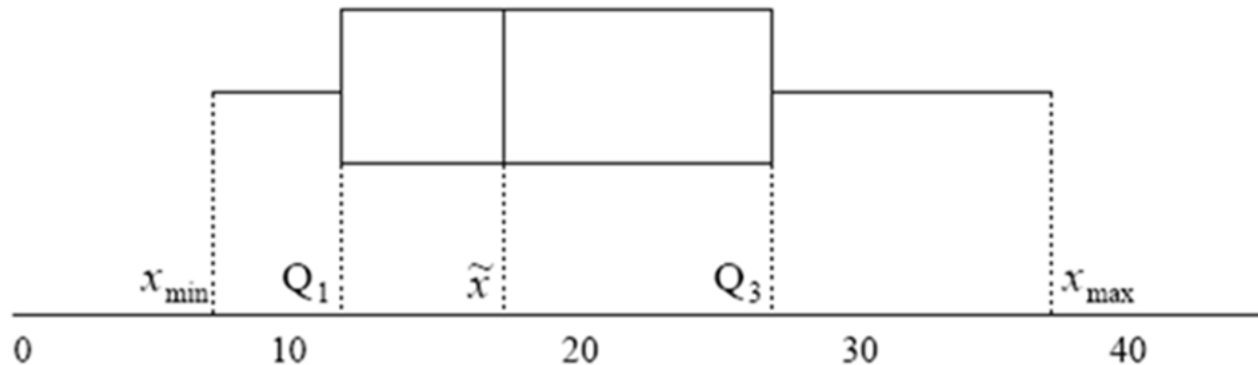
Λογαριθμικοκανονική κατανομή (Log-Normal distribution)

Σε πολλές περιπτώσεις τα αποτελέσματα που λαμβάνονται δεν έχουν κανονική κατανομή.



Η καμπύλη του Σχήματος δείχνει μια στρέβλωση με την επικρατούσα (mode), τη διάμεσο (median) και τη μέση τιμή (mean) να είναι εμφανώς διαφορετικές. Για την απόκτηση μιας σχεδόν κανονικής κατανομής υπολογίζονται οι λογάριθμοι και στη συνέχεια τα $\bar{x}$ και s . (Οι αντιλογάριθμοι των δύο αυτών ποσοτήτων αντιπροσωπεύουν τη γεωμετρική μέση τιμή $\bar{x}_g$ και τη γεωμετρική τυπική απόκλιση s_g .)

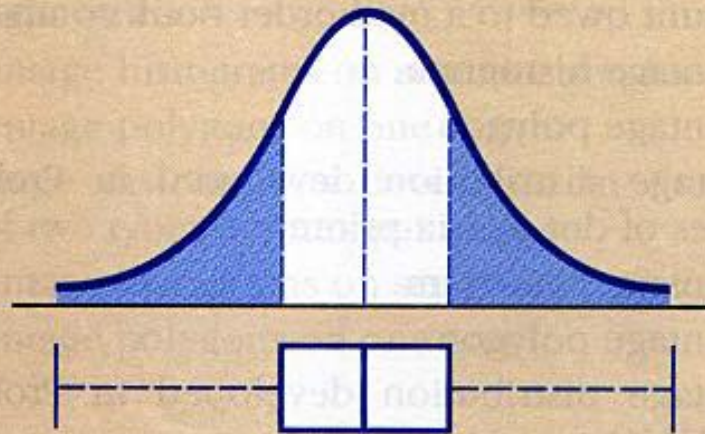
Σύνοψη των 5 αριθμών - **Θηκόγραμμα**



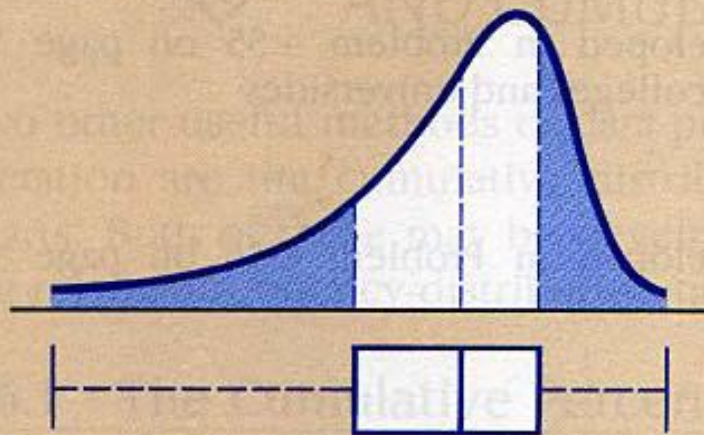
Συνθήκες για αποδοχή κανονικής κατανομής από το Θηκόγραμμα:

- $\tilde{x}$ όχι κοντά στο Q_1 ή στο Q_3
- το εύρος των τιμών στα δύο ακραία τεταρτημόρια να μη διαφέρει σημαντικά (συγκρίσιμα μήκη απολήξεων)
- να μην υπάρχουν ακραίες τιμές

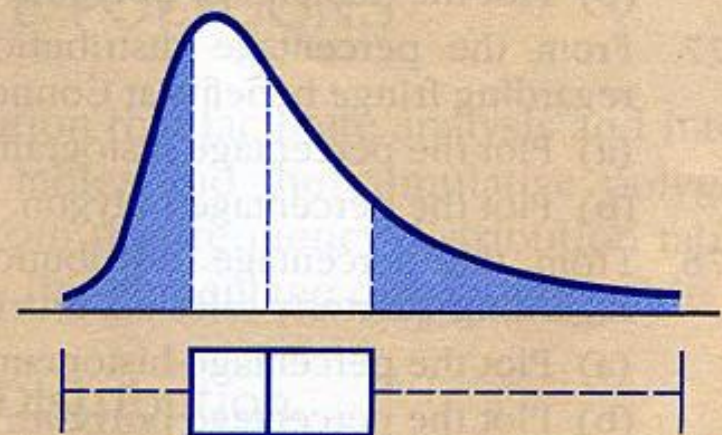
Αν το Θηκόγραμμα δίνει ενδείξεις σημαντικής απόκλισης από τη συμμετρική κατανομή τότε **δεν θα πρέπει** να θεωρήσουμε πως η κατανομή είναι κανονική στη στατιστική ανάλυση που θα κάνουμε στη συνέχεια.



(a) Bell-shaped distribution



(b) Left-skewed distribution



(c) Right-skewed distribution

ΚΑΤΑΝΟΜΗ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ

Αποτελέσματα από 50 προσδιορισμούς συγκέντρωσης νιτρικών ιόντων, σε $\mu\text{g mL}^{-1}$

0,51	0,51	0,51	0,50	0,51	0,49	0,52	0,53	0,50	0,47
0,51	0,52	0,53	0,48	0,49	0,50	0,52	0,49	0,49	0,50
0,49	0,48	0,46	0,49	0,49	0,48	0,49	0,49	0,51	0,47
0,51	0,51	0,51	0,48	0,50	0,47	0,50	0,51	0,49	0,48
0,51	0,50	0,50	0,53	0,52	0,52	0,50	0,50	0,51	0,51

Πίνακας συχνοτήτων για μετρήσεις της συγκέντρωσης νιτρικών ιόντων

Συγκέντρωση νιτρικών ιόντων ($\mu\text{g mL}^{-1}$)

Συχνότητα

0,46

1

0,47

3

0,48

5

0,49

10

0,50

10

0,51

13

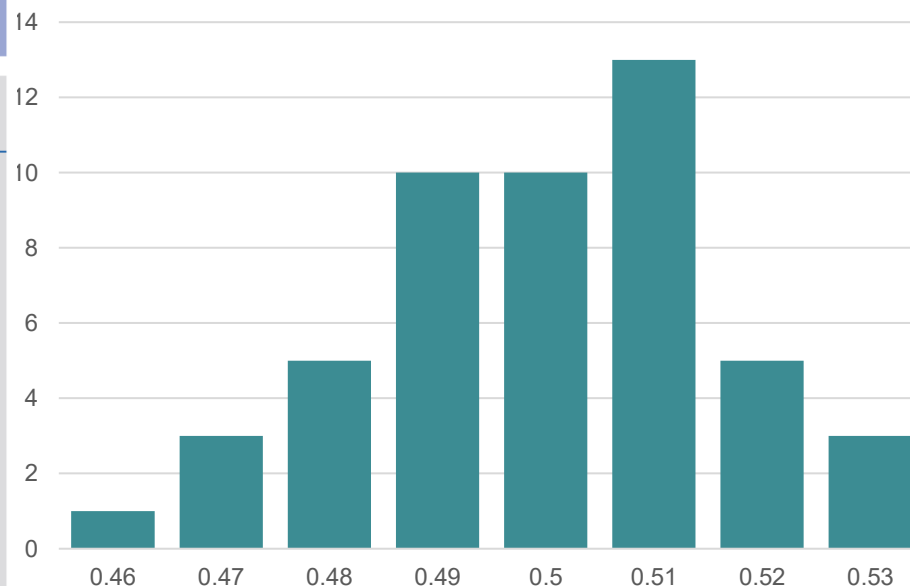
0,52

5

0,53

3

Μέση τιμή: 0,50



ΚΑΤΑΝΟΜΗ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ

Πίνακας συχνοτήτων για μετρήσεις της συγκέντρωσης νιτρικών ιόντων

Συγκέντρωση νιτρικών ιόντων ($\mu\text{g mL}^{-1}$)	Συχνότητα
0,46	1
0,47	3
0,48	5
0,49	10
0,50	10
0,51	13
0,52	5
0,53	3

Μέση τιμή: 0,50

Πρακτικά, είναι ασυνήθιστο να γίνουν 50 επαναλαμβανόμενες μετρήσεις. Ένας πιθανότερος αριθμός θα ήταν 5(?).

Θεωρούμε κάθε στήλη ως **ένα δείγμα** και τα αποτελέσματα του πίνακα ως **10 δείγματα** (καθένα περιέχει 5 μετρήσεις).

Αποτελέσματα από 50 προσδιορισμούς συγκέντρωσης νιτρικών ιόντων, σε $\mu\text{g mL}^{-1}$

0,51	0,51	0,51	0,50	0,51	0,49	0,52	0,53	0,50	0,47
0,51	0,52	0,53	0,48	0,49	0,50	0,52	0,49	0,49	0,50
0,49	0,48	0,46	0,49	0,49	0,48	0,49	0,49	0,51	0,47
0,51	0,51	0,51	0,48	0,50	0,47	0,50	0,51	0,49	0,48
0,51	0,50	0,50	0,53	0,52	0,52	0,50	0,50	0,51	0,51

ΚΑΤΑΝΟΜΗ ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ

Πίνακας συχνοτήτων για μετρήσεις της συγκέντρωσης νιτρικών ιόντων

Συγκέντρωση νιτρικών ιόντων ($\mu\text{g mL}^{-1}$)	Συχνότητα
0,46	1
0,47	3
0,48	5
0,49	10
0,50	10
0,51	13
0,52	5
0,53	3

Μέση τιμή: 0,50

Οι μέσες τιμές αυτών είναι: 0,506, 0,504, 0,502, 0,496, 0,502, 0,492, 0,506, 0,504, 0,500 και 0,486.

Τι παρατηρούμε

I. οι μέσες τιμές αυτών των δειγμάτων διασπείρονται γύρω από τη μέση τιμή.

II. οι μέσες τιμές είναι περισσότερο συγκεντρωμένες από τις αρχικές μετρήσεις.

Η κατανομή όλων των πιθανών μέσων τιμών των δειγμάτων (στην περίπτωση του πληθυσμού ένας άπειρος αριθμός) καλείται **δειγματική κατανομή της μέσης τιμής**.

Η μέση τιμή είναι ίδια με αυτή του αρχικού πληθυσμού. Η τυπική απόκλιση εδώ, καλείται **τυπικό σφάλμα της μέσης τιμής**. Για ένα δείγμα n μετρήσεων ισχύει:

Τυπικό σφάλμα μέσης τιμής (standard error of the mean) = $\sigma / \sqrt{n}$

σ είναι η τυπική απόκλιση του πληθυσμού

Θεώρημα Κεντρικού Ορίου (Central Limit Theorem)

Περιγράφει τα χαρακτηριστικά του "πληθυσμού των μέσων τιμών" που σχηματίζεται από τις μέσες τιμές άπειρων πληθυσμιακών δειγμάτων καθένα από τα οποία αποτελείται από n τυχαία δεδομένα (n είναι το μέγεθος του πληθυσμιακού δείγματος), τα οποία έχουν ληφθεί από τον ίδιο "μητρικό πληθυσμό".

Το Θεώρημα αυτό προβλέπει ότι ανεξάρτητα από τη μορφή της κατανομής του μητρικού πληθυσμού:

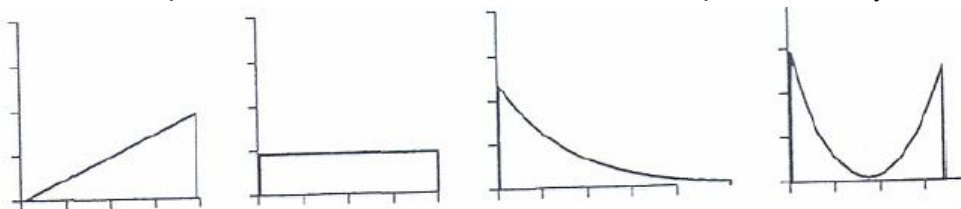
- I.** Η μέση τιμή του νέου πληθυσμού των μέσων τιμών είναι πάντοτε ίση προς τη μέση τιμή του μητρικού πληθυσμού από τον οποίο ελήφθησαν τα δείγματα.
- II.** Η τυπική απόκλιση του πληθυσμού των μέσων τιμών είναι πάντοτε ίση προς την τυπική απόκλιση του μητρικού πληθυσμού σ διαιρεμένη με την τετραγωνική ρίζα του μεγέθους του πληθυσμιακού δείγματος n .
- III.** [**Και το πιο σημαντικό**] Η κατανομή του πληθυσμού των μέσων τιμών τείνει προς την κανονική (κατά Gauss) κατανομή, όσο αυξάνει το μέγεθος n του πληθυσμιακού δείγματος.

Συνεπώς: αν λαμβάνουμε τις μέσες τιμές μετρήσεων μιας ορισμένης ποσότητας, η κατανομή τους τείνει προς την "κανονική". Επιπλέον, το τυχαίο σφάλμα που συνοδεύει τις μετρήσεις μας τείνει, επίσης, ν' αποκτήσει κανονική κατανομή.

Το Θεώρημα, ερμηνεύει συχνά την εμφάνιση στις μετρητικές διαδικασίες της "κανονικής κατανομής" ή "κατανομής κατά Gauss" (που είναι γνωστή και ως σχήματος κώδωνα).

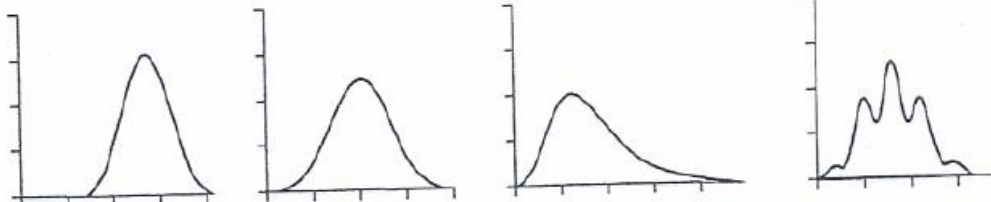
Θεώρημα Κεντρικού Ορίου (Central Limit Theorem)

Ας θεωρήσουμε τις παρακάτω κατανομές τεσσάρων πληθυσμών

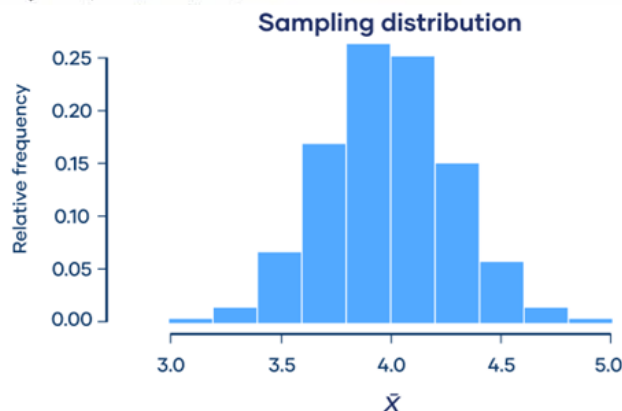
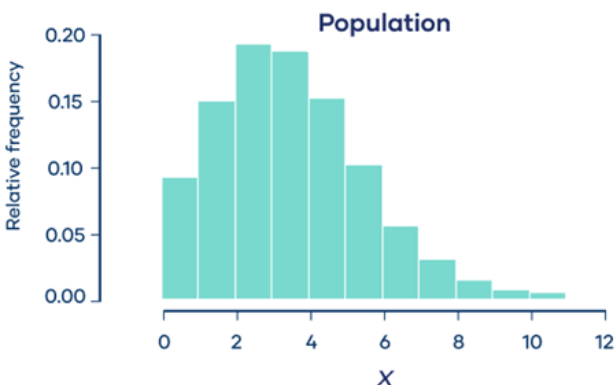
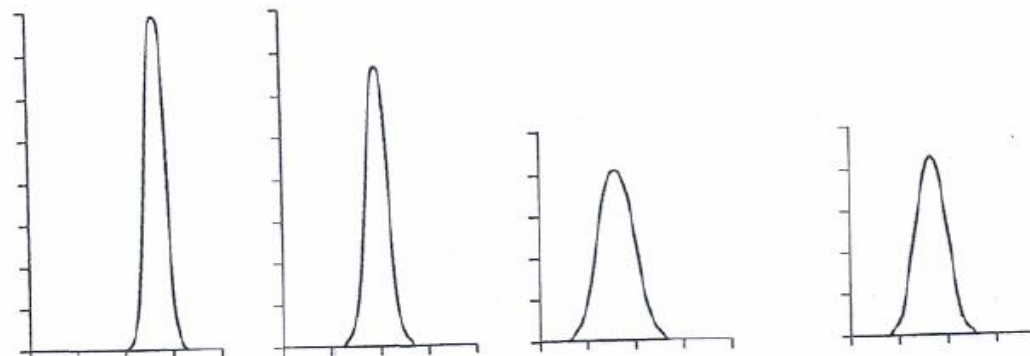


τότε, οι κατανομές των δειγματικών μέσων είναι αντίστοιχα,

για $n = 4$



για $n = 25$



Αν λάβουμε π.χ. 10.000 δείγματα από έναν πληθυσμό με $n=50$ οι μέσες τιμές ακολουθούν κανονική κατανομή, σύμφωνα με το Θεώρημα κεντρικού ορίου.

Θεώρημα Κεντρικού Ορίου (Central Limit Theorem)

- Όσο μεγαλύτερο είναι το μέγεθος n των δειγμάτων τόσο ακριβέστερη είναι η προσέγγιση της κατανομής των δειγματικών μέσων τιμών από την κανονική κατανομή.
 - Το τυπικό σφάλμα είναι μικρότερο από την τυπική απόκλιση σ της τ.μ. X .
Δηλαδή, η μεταβλητότητα (διακύμανση) των δειγματικών μέσων είναι μικρότερη από τη μεταβλητότητα του πληθυσμού από τον οποίο προέρχονται.
- Όσο αυξάνει το μέγεθος του δείγματος, η δειγματική μεταβλητότητα (η μεταβλητότητα από δείγμα σε δείγμα των μέσων τιμών) ελαττώνεται.

Βαθμοί ελευθερίας (degrees of freedom, df)

Όρος που αναφέρεται στο ελάχιστο σύνολο μεταβλητών που απαιτούνται για τον προσδιορισμό της κατάστασης ενός συστήματος ή
το ελάχιστο των στοιχείων που απαιτούνται για την αναπαραγωγή της πληροφορίας που περιέχεται σε ένα σύνολο ή
ο αριθμός των ανεξάρτητων μετρήσεων-παρατηρήσεων που απαιτούνται για τον υπολογισμό ενός **στατιστικού στοιχείου**.

Οι βαθμοί ελευθερίας εξαρτώνται από τους **περιορισμούς** που επιβάλλουμε και συγκεκριμένα ελαττώνονται όσο περισσότερους περιορισμούς επιβάλλουμε γιατί σ' αυτούς μεταφέρεται ένα μέρος της συνολικής πληροφορίας.

Βαθμοί ελευθερίας στη γεωμετρία και τη φυσική: Σημείο στο χώρο, περιγράφεται από τρεις βαθμούς ελευθερίας. Η κίνηση στο επίπεδο από δύο βαθμούς ελευθερίας και στην ευθεία από ένα βαθμό ελευθερίας.

Βαθμοί ελευθερίας μιας στατιστικής κατανομής: ορίζεται ο αριθμός των ανεξάρτητων παρατηρήσεων-μετρήσεων N ελαττωμένος τόσο όσος ο αριθμός των παραμέτρων της κατανομής που υπολογίζονται από τις παρατηρήσεις.
Σε ένα πληθυσμό N μετρήσεων που κατανέμονται κανονικά, τα στατιστικά μ και σ περιγράφονται από τις N μετρήσεις. Άρα, N είναι οι βαθμοί ελευθερίας.

Βαθμοί ελευθερίας (degrees of freedom)

Δηλαδή, έστω ένα σύνολο N αριθμών $\{x_1, x_2, \dots, x_N\}$. Η πληροφορία του συνόλου περιέχεται σε αυτούς τους αριθμούς, άρα έχουμε N βαθμούς ελευθερίας ($df=N$). Αν η μέση τιμή του συνόλου έχει μία σταθερή τιμή μ , τότε αρκεί να προσδιορίσουμε οποιουσδήποτε $N-1$ αριθμούς για να προσδιοριστεί κι αυτός που απομένει, π.χ.

$$x_1 = N\mu - \sum_{i=2}^N x_i$$

Τότε, έχουμε $N-1$ βαθμούς ελευθερίας, $df=N-1$.

Για τον υπολογισμό της τυπικής απόκλισης s μιας σειράς δεδομένων διαιρούμε το άθροισμα των τετραγώνων των αποκλίσεων με το $N-1$. Ο όρος $N-1$ αντιπροσωπεύει τους βαθμούς ελευθερίας.

Σε μια κατανομή στατιστικού δείγματος ισχύει

$$\sum (\bar{x} - x_i) = 0$$

Γνωρίζοντας τις $N-1$ διαφορές, η N -οστή είναι αυστηρά καθορισμένη.

Διάστημα (όρια) εμπιστοσύνης - Επίπεδο σημαντικότητας P

Οι μετρήσεις μας έχουν πάντα μια διασπορά που εξαρτάται από τα χαρακτηριστικά του πειράματος, δηλ. τους τυχαίους παράγοντες (π.χ. σφάλματα) που υπεισέρχονται και από τον τρόπο που έχει σχεδιαστεί ή/και εκτελεστεί το πείραμα.

Η διασπορά των τιμών σε συγκεκριμένα διαστήματα εκφράζεται από την κατανομή πιθανοτήτων.

Μπορούμε να βρούμε τις μέσες τιμές και διασπορές των επιμέρους δειγμάτων και να τις χρησιμοποιήσουμε για να ανακαλύψουμε τα βασικά χαρακτηριστικά του πληθυσμού.

Για τις μετρήσεις μας και τα στατιστικά της κατανομής τους, μπορούμε να εκφράσουμε ποσοτικά την **αβεβαιότητα** που χαρακτηρίζει την εκτίμησή μας με τη μορφή, συνήθως, της **τυπικής απόκλισης** $\bar{x} \pm s$

Για τη μέση τιμή λέμε, "από την τιμή A μέχρι την B , με $X\%$ **επίπεδο εμπιστοσύνης**"

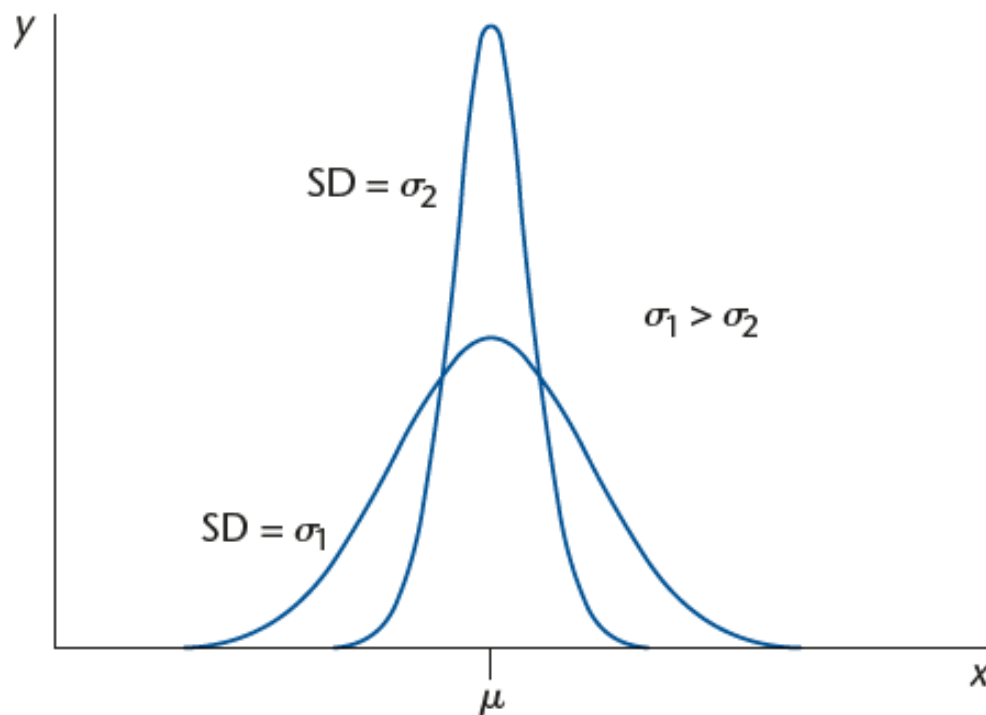
Μεγαλύτερο το X (κοντά στο 100%), μεγαλύτερο το **διάστημα εμπιστοσύνης** ($A-B$).
Μικρότερο το $X(\%)$, στενότερο το διάστημα γύρω από μια προβλεπόμενη τιμή.

Διάστημα (όρια) εμπιστοσύνης - Επίπεδο σημαντικότητας p (ή P)

Η εκτίμησή μας για τη μέση τιμή ενός πληθυσμού, βάσει των δειγμάτων που παίρνουμε, ακολουθεί την κανονική κατανομή σύμφωνα με το Κεντρικό Οριακό Θεώρημα.

Όσο πιο στενή είναι η κορυφή της κατανομής τόσο πιο βέβαιοι είμαστε για το πού βρίσκεται η πραγματική μέση τιμή.

Όσο πιο πλατιά, τόσο μεγαλύτερη η πιθανότητα να βρίσκεται πιο μακριά από την κορυφή της καμπύλης.



Διάστημα (όρια) εμπιστοσύνης - Επίπεδο σημαντικότητας p

Τα όρια A και B του διαστήματος εμπιστοσύνης προκύπτουν προσθέτοντας και αφαιρώντας από την προβλεπόμενη τιμή, μια τιμή που δείχνει το εύρος της αβεβαιότητάς μας.

Τα πιο συνήθη, στην πράξη, **επίπεδα εμπιστοσύνης ή αξιοπιστίας** είναι 95%, 99% ή 99,5%.

Όταν το επίπεδο εμπιστοσύνης αναφέρεται στη διαφορά μίας παραμέτρου ή μίας κατανομής αποτελεσμάτων από κάποια τιμή ή σύνολο τιμών αναφοράς, λέγεται **επίπεδο σημαντικότητας p** .

Η σημαντικότητα μας δείχνει αν η διαφορά είναι υπαρκτή και πρέπει να ληφθεί υπ' όψιν ως τέτοια ή θα μπορούσε να προκύψει π.χ. κατά τύχη ή από άλλες αιτίες που είναι γνωστές αλλά δε μας ενδιαφέρουν.

Μέγεθος δείγματος

Η διαδικασία που περιγράψαμε μέχρι τώρα προϋποθέτει αρκετά μεγάλα δείγματα ($n > 30$) για να έχει εφαρμογή το Θεώρημα Κεντρικού Ορίου

Αν το σ είναι άγνωστο, χρησιμοποιείται ως εκτίμηση το s .

Αν το δείγμα είναι μικρό (π.χ. $n < 20$), τότε, αν γνωρίζω, περιμένω ή θεωρώ λογικό ότι ακολουθεί κανονική κατανομή, χρησιμοποιώ την κατανομή t κατά Student και το στατιστικό στοιχείο s για την τυπική απόκλιση.

Η κατανομή t κατά Student μοιάζει με την κανονική κατανομή, είναι κι αυτή συμμετρική και μάλιστα για μεγάλα δείγματα τείνει να ταυτιστεί με την κανονική.

Για την περίπτωση χρήσης της κατανομής κατά Student, το διάστημα εμπιστοσύνης για τη μέση τιμή είναι:

$$\bar{x} \pm t_{n-1} s / \sqrt{n}$$

κρίσιμη τιμή t

Πίνακας Η κατανομή t

Τιμή t για ένα διάστημα εμπιστοσύνης της κρίσιμης τιμής $ t $ για τιμές P των βαθμών ελευθερίας	90% 0,10	95% 0,05	98% 0,02	99% 0,01
1	6,31	12,71	31,82	63,66
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
12	1,78	2,18	2,68	3,05
14	1,76	2,14	2,62	2,98
16	1,75	2,12	2,58	2,92
18	1,73	2,10	2,55	2,88
20	1,72	2,09	2,53	2,85
30	1,70	2,04	2,46	2,75
50	1,68	2,01	2,40	2,68
∞	1,64	1,96	2,33	2,58

Οι κρίσιμες τιμές είναι κατάλληλες για αμφίπλευρες δοκιμασίες. Για μονόπλευρες δοκιμασίες, η τιμή λαμβάνεται από τη στήλη που είναι διπλάσια της επιθυμητής τιμής P , π.χ. για μία μονόπλευρη δοκιμασία, $P = 0,05$ και 5 βαθμούς ελευθερίας, η κρίσιμη τιμή λαμβάνεται από τη στήλη για $P = 0,10$ και είναι ίση με 2,02.

Όρια εμπιστοσύνης της μέσης τιμής για μικρά δείγματα

Επαναλαμβάνω: όσο μικρότερο είναι το μέγεθος του δείγματος, τόσο η τυπική απόκλιση s γίνεται λιγότερο αξιόπιστο εκτίμημα του σ .

Για μικρά δείγματα, τα όρια εμπιστοσύνης της μέσης τιμής είναι

$$\bar{x} \pm t_{n-1} s / \sqrt{n}$$

Παράδειγμα

Η περιεκτικότητα νατρίου σε δείγμα ούρων με εκλεκτικό ηλεκτρόδιο ιόντων βρέθηκε ότι είναι 102, 97, 99, 98, 101, 106 mM μετά από σχετικό αριθμό μετρήσεων. Ποια είναι τα όρια εμπιστοσύνης της συγκέντρωσης νατρίου για επίπεδα εμπιστοσύνης 95% και 99% ?

Η μέση τιμή και η τυπική απόκλιση είναι 100,5 mM και 3,27 mM αντίστοιχα. Υπάρχουν έξι μετρήσεις και 5 βαθμοί ελευθερίας. Από τον πίνακα των τιμών t για επίπεδα εμπιστοσύνης 95% και 99%, το $t(5)$ είναι 2,57 και 4,03 αντίστοιχα. Άρα

$$100,5 \pm 2,57 \times 3,27 / \sqrt{6} = 100,5 \pm 3,4 \text{ mM}$$

$$100,5 \pm 4,03 \times 3,27 / \sqrt{6} = 100,5 \pm 5,4 \text{ mM}$$

Τα όρια εμπιστοσύνης μπορούν να χρησιμοποιηθούν και για τη δοκιμασία ύπαρξης συστηματικού σφάλματος.

Παράδειγμα

Η κλίμακα απορρόφησης ενός φωτομέτρου δοκιμάζεται σε ένα συγκεκριμένο μήκος κύματος με πρότυπο διάλυμα με δεδομένη απορρόφηση 0,470. Δέκα μετρήσεις της απορρόφησης με το συγκεκριμένο φωτόμετρο δίνουν $\bar{x} = 0,461$ και $s = 0,003$.

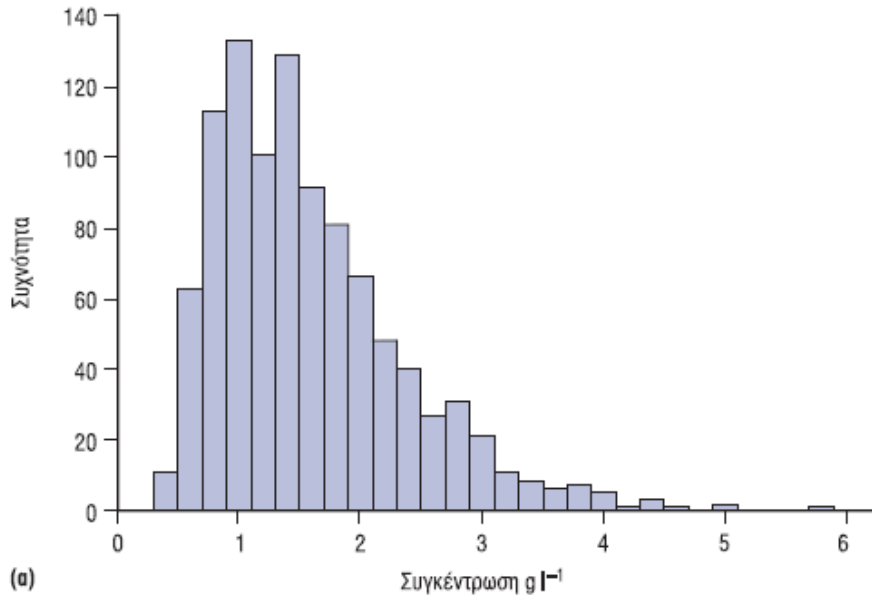
Να βρεθούν τα όρια εμπιστοσύνης (95%) για τη μέση απορρόφηση και να αποφασιστεί αν υπάρχει συστηματικό σφάλμα στη μέτρηση με το φωτόμετρο.

Το διάστημα εμπιστοσύνης, για επίπεδο εμπιστοσύνης (ή αξιοπιστίας) 95%, όπως μετράται με το φωτόμετρο, είναι:

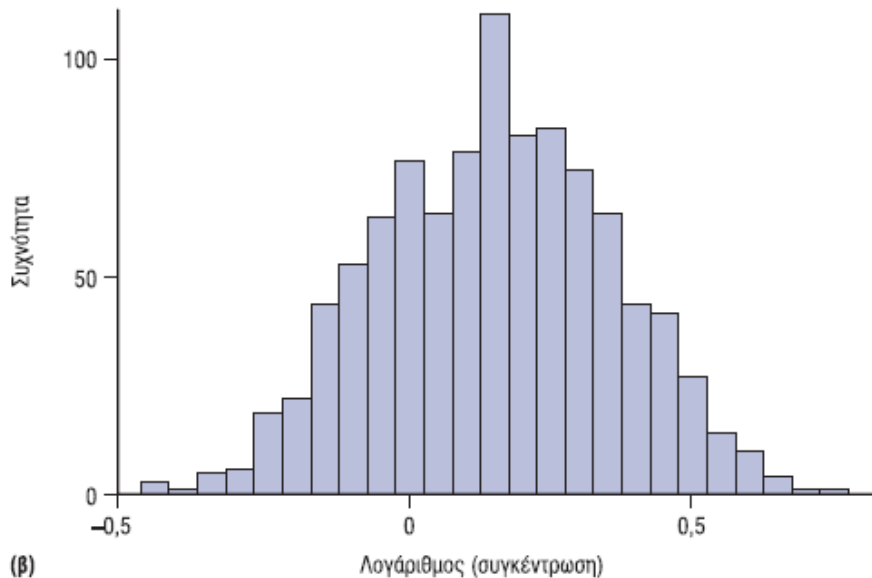
$$\bar{x} \pm t_{n-1} s / \sqrt{n} = 0,461 \pm 2,26 \times 0,003 / \sqrt{10} = 0,461 \pm 0,002$$

Καθώς το διάστημα εμπιστοσύνης δεν περιλαμβάνει τη σωστή απορρόφηση 0,470, είναι πιθανό να υπάρχει συστηματικό σφάλμα.

Όρια εμπιστοσύνης της γεωμετρικής μέσης τιμής για λογαριθμικοκανονική κατανομή



Επειδή οι λογάριθμοι των μετρήσεων παρουσιάζουν κανονική κατανομή, είναι πιο ακριβές να εργαστεί κανείς με τους λογαρίθμους των μετρήσεων.



Παράδειγμα

Οι επόμενες τιμές (%) δίνουν τη συγκέντρωση ενός αντισώματος στον ορό του ανθρώπινου αίματος ο οποίος ελήφθη από οκτώ υγιείς ενήλικες.

2,15 1,13 2,04 1,45 1,35 1,09 0,99 2,07

Να υπολογιστεί το διάστημα εμπιστοσύνης (για επίπεδο εμπιστοσύνης 95%) υποθέτοντας ότι η συγκέντρωση του αντισώματος παρουσιάζει λογαριθμικοκανονική κατανομή.

Οι λογάριθμοι (με βάση το 10) είναι:

0,332 0,053 0,310 0,161 0,130 0,037 -0,004
0,316

Η μέση τιμή των λογαρίθμων είναι 0,1669 και η τυπική απόκλιση είναι 0,1365 (δίνοντας $10^{0,1669} = 1,47$ (%), ως πραγματική μέση τιμή).

Τα όρια αξιοπιστίας για επίπεδο εμπιστοσύνης 95% των λογαριθμημένων τιμών είναι:

$$0,1669 \pm 2,36 \times 0,1365 / \sqrt{8} = 0,1669 \pm 0,1139 = 0,0530 \text{ έως } 0,2808$$

Παίρνοντας τους αντιλογάριθμους των ορίων αυτών, τα όρια εμπιστοσύνης της μέσης τιμής προκύπτουν ως **1,13%** και **1,91%**.

ΣΥΜΠΕΡΑΣΜΑΤΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Δοκιμασίες (ή δοκιμές ή έλεγχοι) σημαντικότητας (significance tests)

Η έννοια της στατιστικά σημαντικής διαφοράς (statistically significant difference):

Ο σημαντικός δεν έχει σχέση με μεγέθη, δηλ. δεν σημαίνει μεγάλος, υπερβολικός. Σημαντική διαφορά σημαίνει διαφορά οφειλόμενη σε συστηματικά αίτια, σε αντίθεση με τη μη σημαντική διαφορά που οφείλεται σε τυχαία αίτια.

Από τις πιο σημαντικές ιδιότητες μιας αναλυτικής μεθόδου είναι η έλλειψη συστηματικών σφαλμάτων. Η ιδιότητα αυτή μπορεί να εκτιμηθεί με την ανάλυση (μέρους) ενός δείγματος με γνωστή συγκέντρωση αναλύτη. Ωστόσο, και το τυχαίο σφάλμα μπορεί να δώσει τιμές που αποκλίνουν από την πραγματική.

Για να αποφανθούμε αν η διαφορά της μετρούμενης από την πραγματική τιμή οφείλεται σε τυχαία σφάλματα, εφαρμόζουμε μια δοκιμασία σημαντικότητας.

Σκοπός, δηλαδή, είναι να επιβεβαιωθεί αν η διαφορά μεταξύ δύο αποτελεσμάτων είναι στατιστικά σημαντική ή οφείλεται σε τυχαίες διακυμάνσεις.

1. Σύγκριση μιας μέσης τιμής με μια γνωστή-πραγματική τιμή
2. Σύγκριση μεταξύ δύο μέσων τιμών
3. Δοκιμασία t για σύγκριση ζευγών τιμών (μονόπλευρη, αμφίπλευρη)
4. Δοκιμασία F για σύγκριση τυπικών αποκλίσεων

Μηδενική υπόθεση

Σύγκριση αναλυτικών αποτελεσμάτων που ελήφθησαν:

Περίπτωση 1. με την ίδια μέθοδο στα δείγματα Α και Β, για να διαπιστωθεί αν τα δείγματα περιέχουν τη μετρούμενη ουσία σε ίδια ή διαφορετική συγκέντρωση.

Περίπτωση 2. με δύο διαφορετικές μεθόδους Α και Β στο ίδιο δείγμα, για να διαπιστωθεί αν οι δύο μέθοδοι παρέχουν ίδια ή διαφορετικά αποτελέσματα.

Το αποτέλεσμα αυτών των δοκιμασιών είναι η αποδοχή ή η απόρριψη της μηδενικής υπόθεσης (null hypothesis, H_0). Η μηδενική υπόθεση, γενικά, δηλώνει ότι: "Οι διαφορές - αποκλίσεις οφείλονται αποκλειστικά σε τυχαία και όχι σε συστηματικά αίτια-σφάλματα". Η εναλλακτική υπόθεση (alternative hypothesis, H_a) δηλώνει ακριβώς το αντίθετο.

Ποια είναι η μηδενική υπόθεση για τις παραπάνω περιπτώσεις?

Οι μέσες τιμές είναι ίδιες, δηλ. για την Περίπτωση 1: η περιεκτικότητα των δύο δειγμάτων σε προσδιοριζόμενη ουσία είναι η ίδια, για την Περίπτωση 2: και οι δύο μέθοδοι παρέχουν τα ίδια αναλυτικά αποτελέσματα. Οι παρατηρούμενες διαφορές (αν υπάρχουν) οφείλονται καθαρά σε τυχαία σφάλματα.

Εναλλακτική υπόθεση: Οι μέσες τιμές διαφέρουν σημαντικά, δηλ. για την Περίπτωση 1: οι περιεκτικότητες των δύο δειγμάτων σε προσδιοριζόμενη ουσία είναι διαφορετικές, για την Περίπτωση 2: οι μέθοδοι παρέχουν διαφορετικά αποτελέσματα (επομένως τουλάχιστον η μία παρουσιάζει συστηματικό αναλυτικό σφάλμα).

1. Σύγκριση μιας μέσης τιμής με μια γνωστή τιμή

Η μηδενική υπόθεση (null hypothesis, H_0): Δεν υπάρχει διαφορά μεταξύ των δύο συγκρίσιμων τιμών παρά μόνο εξ αιτίας τυχαίων διακυμάνσεων.

Υποθέτοντας ότι η μηδενική υπόθεση ισχύει, η πιθανότητα ότι η παρατηρούμενη διαφορά μεταξύ της μέσης τιμής του δείγματος και της γνωστής τιμής προκύπτει ως αποτέλεσμα τυχαίων σφαλμάτων.

Ορισμός επιπέδου αξιοπιστίας

Επίπεδο αξιοπιστίας 95% ή επίπεδο σημαντικότητας $p = 0,05$ (ή 5%): αν η H_0 απορριφθεί, υπάρχει πιθανότητα 5 στις 100 περιπτώσεις να έχει απορριφθεί η μηδενική υπόθεση ενώ ισχύει.

Όσο μικρότερη (μεγαλύτερη) είναι η πιθανότητα η παρατηρούμενη διαφορά να συμβεί τυχαία τόσο λιγότερο (περισσότερο) πιθανό είναι η μηδενική υπόθεση να ισχύει.

Για να είμαστε πιο σίγουροι ότι λαμβάνουμε τη σωστή απόφαση μπορεί να χρησιμοποιηθεί υψηλότερο επίπεδο σημαντικότητας, συνήθως 0,01 ή 0,001 (1% ή 0,1%).

Το επίπεδο σημαντικότητας υποδεικνύεται γράφοντας, για παράδειγμα, p (δηλαδή πιθανότητα) = 0,05 και δίνει την πιθανότητα απόρριψης μιας ισχύουσας μηδενικής υπόθεσης.

Εάν δεν απορριφθεί η μηδενική υπόθεση, δεν έχει αποδειχθεί ότι ισχύει. Μόνο ότι δεν έχει αποδειχθεί ότι είναι ψευδής.

1. Σύγκριση μιας μέσης τιμής με μια γνωστή τιμή

Η μηδενική υπόθεση (null hypothesis, H_0): Δεν υπάρχει διαφορά μεταξύ των δύο συγκρίσιμων τιμών παρά μόνο εξ αιτίας τυχαίων διακυμάνσεων.

Επίπεδο αξιοπιστίας 95% ή επίπεδο σημαντικότητας $p = 0,05$ (ή 5%): αν η H_0 απορριφθεί, υπάρχει πιθανότητα 5 στις 100 περιπτώσεις να έχει απορριφθεί η μηδενική υπόθεση ενώ ισχύει.

Πρέπει να αποφανθούμε αν: μέση τιμή = πραγματική (γνωστή) τιμή - δηλαδή αν ισχύει η μηδενική υπόθεση, υπολογίζεται ο **στατιστικός όρος** ή **στατιστικό στοιχείο (statistic)**:

$$t = (\bar{x} - x_{\text{πραγμ.}}) \sqrt{n} / s$$

Αν $|t| > t_{\text{κρίσιμη}}$ (θεωρητική) τιμή, η H_0 απορρίπτεται (για $n-1$ βαθμούς ελευθερίας).

Παράδειγμα: Προσδιορισμός Se σε νερό πιστοποιημένο υλικό αναφοράς (CRM) συγκέντρωσης 50 ng/mL.

Βρέθηκαν οι συγκεντρώσεις: 50,4 50,7 49,1 49,0 51,1 ng/mL

Υπάρχει ένδειξη συστηματικού σφάλματος?

Μέση τιμή: 50,06 ng/mL

τυπική απόκλιση: 0,95 ng/mL

Μηδενική υπόθεση H_0 : δεν υπάρχει συστηματικό σφάλμα

Υπολογίζεται η $t = 0,14$

Κρίσιμη τιμή για επίπεδο σημαντικότητας $p = 0,05$: $t(4) = 2,78 > 0,14$

Η μηδενική υπόθεση ισχύει. Δεν υπάρχει ένδειξη συστηματικού σφάλματος.

Πίνακας Η κατανομή t

Τιμή t για ένα διάστημα εμπιστοσύνης της κρίσιμης τιμής $ t $ για τιμές P των βαθμών ελευθερίας	90% 0,10	95% 0,05	98% 0,02	99% 0,01
1	6,31	12,71	31,82	63,66
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
12	1,78	2,18	2,68	3,05
14	1,76	2,14	2,62	2,98
16	1,75	2,12	2,58	2,92
18	1,73	2,10	2,55	2,88
20	1,72	2,09	2,53	2,85
30	1,70	2,04	2,46	2,75
50	1,68	2,01	2,40	2,68
∞	1,64	1,96	2,33	2,58

Οι κρίσιμες τιμές είναι κατάλληλες για αμφίπλευρες δοκιμασίες. Για μονόπλευρες δοκιμασίες, η τιμή λαμβάνεται από τη στήλη που είναι διπλάσια της επιθυμητής τιμής P , π.χ. για μία μονόπλευρη δοκιμασία, $P = 0,05$ και 5 βαθμούς ελευθερίας, η κρίσιμη τιμή λαμβάνεται από τη στήλη για $P = 0,10$ και είναι ίση με 2,02.

2. Σύγκριση μεταξύ δύο μέσων τιμών

Μια αναλυτική μέθοδος δοκιμάζεται με σύγκρισή της με μια δεύτερη μέθοδο (πιθανόν μέθοδος αναφοράς).

H_0 : Οι δύο μέθοδοι δίνουν τα ίδια αποτελέσματα. Δηλαδή $\bar{x}_1 = \bar{x}_2$.

Απαιτείται να ελεγχθεί αν η διαφορά των μέσων τιμών διαφέρει σημαντικά από το μηδέν.

Αν οι τυπικές αποκλίσεις των δύο μεθόδων δεν είναι σημαντικά διαφορετικές, τότε η συνολική τυπική απόκλιση s και ο στατιστικός όρος t , υπολογίζονται

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}$$

$n_1 + n_2 - 2$, βαθμοί ελευθερίας

Παράδειγμα: Συγκρίνονται δύο μέθοδοι για τον προσδιορισμό Cr σε φυτικό ιστό. Πραγματοποιήθηκαν πέντε προσδιορισμοί για κάθε μέθοδο.

Μέθοδος 1: μέση τιμή 1,48 ng/g, τυπική απόκλιση 0,28 ng/g

Μέθοδος 2: μέση τιμή 2,33 ng/g, τυπική απόκλιση 0,31 ng/g

Μηδενική υπόθεση: Οι μέσες τιμές των δύο μεθόδων είναι στατιστικά ίσες.

Συνολική τυπική απόκλιση $s = 0,29$ και $t = 4,56$.

Για 8 βαθμούς ελευθερίας η κρίσιμη τιμή $t(8) = 2,31$ (για $P = 0,05$). Η μηδενική υπόθεση απορρίπτεται.

ΟΤΑΝ ΤΑ ΔΕΙΓΜΑΤΑ ΔΕΝ ΠΡΟΕΡΧΟΝΤΑΙ ΑΠΟ ΤΟΝ ΙΔΙΟ ΠΛΗΘΥΣΜΟ (ΔΗΛ. ΔΙΑΦΟΡΕΤΙΚΕΣ ΤΥΠΙΚΕΣ ΑΠΟΚΛΙΣΕΙΣ) ΥΠΑΡΧΟΥΝ ΑΛΛΟΙ ΤΡΟΠΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΚΑΙ ΣΥΓΚΡΙΣΗΣ

Πίνακας Η κατανομή t

Τιμή t για ένα διάστημα εμπιστοσύνης της κρίσιμης τιμής $ t $ για τιμές P των βαθμών ελευθερίας	90% 0,10	95% 0,05	98% 0,02	99% 0,01
1	6,31	12,71	31,82	63,66
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
12	1,78	2,18	2,68	3,05
14	1,76	2,14	2,62	2,98
16	1,75	2,12	2,58	2,92
18	1,73	2,10	2,55	2,88
20	1,72	2,09	2,53	2,85
30	1,70	2,04	2,46	2,75
50	1,68	2,01	2,40	2,68
∞	1,64	1,96	2,33	2,58

Οι κρίσιμες τιμές είναι κατάλληλες για αμφίπλευρες δοκιμασίες. Για μονόπλευρες δοκιμασίες, η τιμή λαμβάνεται από τη στήλη που είναι διπλάσια της επιθυμητής τιμής P , π.χ. για μία μονόπλευρη δοκιμασία, $P = 0,05$ και 5 βαθμούς ελευθερίας, η κρίσιμη τιμή λαμβάνεται από τη στήλη για $P = 0,10$ και είναι ίση με 2,02.

3. Δοκιμασία t για σύγκριση ζευγών (μονόπλευρη ή αμφίπλευρη)

Συχνά δύο μέθοδοι συγκρίνονται μετρώντας δείγματα με διαφορετικές ποσότητες-συγκεντρώσεις ενός αναλύτη. Επιθυμούμε να γνωρίζουμε αν οι μέθοδοι παράγουν στατιστικά σημαντικά διαφορετικά αποτελέσματα.

Εδώ η προηγούμενη μέθοδος σύγκρισης των μέσων τιμών δεν είναι κατάλληλη διότι δεν διαχωρίζει τις διακυμάνσεις που προέρχονται από τη μέθοδο, από αυτές που προέρχονται από τα δείγματα (τα δείγματα είναι διαφορετικά!).

Η δυσκολία ξεπερνάται εξετάζοντας τη διαφορά μεταξύ των ζευγών των αποτελεσμάτων από τις δύο μεθόδους.

Υπολογίζεται ο στατιστικός όρος-στατιστικό στοιχείο t για n αποτελέσματα-ζεύγη (d = διαφορά)

$$t = \bar{d} \sqrt{n} / s_d$$

Σύγκριση με την θεωρητική τιμή t για $n - 1$ βαθμούς ελευθερίας.

Πίνακας Η κατανομή t

Τιμή t για ένα διάστημα εμπιστοσύνης της κρίσιμης τιμής $ t $ για τιμές P των βαθμών ελευθερίας	90% 0,10	95% 0,05	98% 0,02	99% 0,01
1	6,31	12,71	31,82	63,66
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
12	1,78	2,18	2,68	3,05
14	1,76	2,14	2,62	2,98
16	1,75	2,12	2,58	2,92
18	1,73	2,10	2,55	2,88
20	1,72	2,09	2,53	2,85
30	1,70	2,04	2,46	2,75
50	1,68	2,01	2,40	2,68
∞	1,64	1,96	2,33	2,58

Οι κρίσιμες τιμές είναι κατάλληλες για αμφίπλευρες δοκιμασίες. Για μονόπλευρες δοκιμασίες, η τιμή λαμβάνεται από τη στήλη που είναι διπλάσια της επιθυμητής τιμής P , π.χ. για μία μονόπλευρη δοκιμασία, $P = 0,05$ και 5 βαθμούς ελευθερίας, η κρίσιμη τιμή λαμβάνεται από τη στήλη για $P = 0,10$ και είναι ίση με 2,02.

Παράδειγμα: Αποτελέσματα προσδιορισμού παρακεταμόλης (mg/g) σε δέκα ταμπλέτες

<u>Δείγμα</u>	<u>UV-μέθοδος</u>	<u>Near IR μέθοδος</u>
1	84,63	83,15
2	84,38	83,72
3	84,08	83,84
4	84,41	84,20
5		
6		
7		
8		
9		
10	84,03	84,24

Οι διαφορές μεταξύ των ζευγών των τιμών είναι:

+1,48, +0,66, +0,24, +0,21,
-0,10, -0,61, -0,10, +0,09,
+0,07, -0,21

$\bar{d} = 0,173$ $s_d = 0,57$ (με τα πρόσημα)

και υπολογίζω $t = 0,88$. Η κρίσιμη τιμή είναι $t(9) = 2,26$ ($P = 0,05$).

Άρα η μηδενική υπόθεση διατηρείται (δηλαδή, οι μέθοδοι δεν δίνουν σημαντικά διαφορετικά αποτελέσματα για τα διαφορετικά δείγματα).

Η δοκιμασία t για σύγκριση ζευγών δεν απαιτεί οι επαναληψιμότητες των δύο μεθόδων να είναι ίσες αλλά οι διαφορές d να προέρχονται από κανονική κατανομή.

Η δοκιμασία t για σύγκριση ζευγών είναι ακατάλληλη για ανάλυση δειγμάτων που οι συγκεντρώσεις τους διαφέρουν πάνω από μια τάξη μεγέθους επειδή βασίζεται στην προϋπόθεση ότι τα τυχαία και συστηματικά σφάλματα είναι ανεξάρτητα της συγκέντρωσης (Απαιτείται ανάλυση γραμμικής συσχέτισης).

Μονόπλευρες ή αμφίπλευρες δοκιμασίες ???

Στα παραδείγματα που περιγράφηκαν ενδιέφερε η ύπαρξη διαφοράς μεταξύ δύο μέσων τιμών προς οποιαδήποτε κατεύθυνση.

Στις περισσότερες περιπτώσεις ο αναλυτής δεν γνωρίζει αν η διαφορά μεταξύ της πειραματικής μέσης τιμής και της τιμής αναφοράς είναι θετική ή αρνητική.

Άρα, η δοκιμασία πρέπει να καλύπτει και τις δύο πιθανότητες (αμφίπλευρη δοκιμασία).

Σε κάποιες άλλες περιπτώσεις, απαιτείται διαφορετικού είδους δοκιμασία (μονόπλευρη).

π.χ. η αύξηση της ταχύτητας μίας αντίδρασης παρουσία καταλύτη απαιτεί δοκιμασία σημαντικότητας προς τη μία (θετική) κατεύθυνση.

Τι γίνεται με τις θεωρητικές τιμές t ?

Για επίπεδο σημαντικότητας $P = 0,05$, η κατάλληλη τιμή για μονόπλευρη δοκιμασία βρίσκεται από τη στήλη που αντιστοιχεί σε $P = 0,10$ του αντίστοιχου πίνακα.

Πίνακας Η κατανομή t

Τιμή t για ένα διάστημα εμπιστοσύνης της κρίσιμης τιμής $ t $ για τιμές P των βαθμών ελευθερίας	90% 0,10	95% 0,05	98% 0,02	99% 0,01
1	6,31	12,71	31,82	63,66
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
12	1,78	2,18	2,68	3,05
14	1,76	2,14	2,62	2,98
16	1,75	2,12	2,58	2,92
18	1,73	2,10	2,55	2,88
20	1,72	2,09	2,53	2,85
30	1,70	2,04	2,46	2,75
50	1,68	2,01	2,40	2,68
∞	1,64	1,96	2,33	2,58

Οι κρίσιμες τιμές είναι κατάλληλες για αμφίπλευρες δοκιμασίες. Για μονόπλευρες δοκιμασίες, η τιμή λαμβάνεται από τη στήλη που είναι διπλάσια της επιθυμητής τιμής P , π.χ. για μία μονόπλευρη δοκιμασία, $P = 0,05$ και 5 βαθμούς ελευθερίας, η κρίσιμη τιμή λαμβάνεται από τη στήλη για $P = 0,10$ και είναι ίση με 2,02.

Παράδειγμα

Υπάρχει η υποψία ότι μια οξεο-βασική ογκομετρική μέθοδος εμπεριέχει σημαντικό σφάλμα λόγω του έγχρωμου δείκτη που χρησιμοποιείται. Συνεπώς, η μέθοδος έχει την τάση να δίνει αποτελέσματα με θετικό συστηματικό σφάλμα. Για να μελετήσουμε την υπόθεση, χρησιμοποιείται διάλυμα του οξέος συγκέντρωσης 0,1 M για την ογκομέτρηση 25,00 mL διαλύματος βάσεως, συγκέντρωσης 0,1 M, με τα ακόλουθα αποτελέσματα.

25,06 25,18 24,87 25,51 25,34 25,41 (mL)

Να γίνει αξιολόγηση των αποτελεσμάτων σχετικά με το αν υπάρχει ένδειξη για θετικό συστηματικό σφάλμα (positive bias).

Για τα δεδομένα ισχύει:

Μέση τιμή: 25,228 mL, τυπική απόκλιση: 0,238 mL

Υιοθετώντας τη μηδενική υπόθεση H_0 , ότι δεν υπάρχει συστηματικό σφάλμα.

$x_{\text{πραγμ}} = 25,00 \text{ mL}$ και

$$t = (25,228 - 25,00) \times \frac{\sqrt{6}}{0,238} = 2,35$$

Η κρίσιμη τιμή είναι $t_5 = 2,02$ ($P = 0,05$, μονόπλευρη).

Η υπολογισθείσα τιμή του t είναι μεγαλύτερη από τη θεωρητική, άρα η μηδενική υπόθεση απορρίπτεται. Συνεπώς, υπάρχει ένδειξη για **θετικό συστηματικό σφάλμα**.

Στο προηγούμενο παράδειγμα

Παρατήρηση: Για μια αμφίπλευρη δοκιμασία ($t_5 = 2,57$), η μηδενική υπόθεση δεν θα απορριπτόταν!

Η απόφαση για εφαρμογή μονόπλευρης ή αμφίπλευρης δοκιμασίας εξαρτάται από την υποψία ή αναμονή θετικού συστηματικού σφάλματος.

Γενικά, οι αμφίπλευρες δοκιμασίες είναι πιο συνηθισμένες απ' τις μονόπλευρες.

Το ίδιο ισχύει και για τη δοκιμασία t για σύγκριση ζευγών.

4. Δοκιμασία F για σύγκριση τυπικών αποκλίσεων

Μέχρι τώρα, οι δοκιμασίες αφορούσαν σύγκριση τιμών για συστηματικά σφάλματα.

Πολύ συχνά, είναι σημαντικό να συγκρίνονται οι τυπικές αποκλίσεις, δηλαδή τα τυχαία σφάλματα σειρών δεδομένων.

Υπάρχουν δύο δυνατότητες σύγκρισης:

- Η μέθοδος A είναι πιο επαναλήψιμη από τη μέθοδο B (μονόπλευρη δοκιμασία).
- Η μέθοδος A και η μέθοδος B διαφέρουν ως προς την επαναληψιμότητα (αμφίπλευρη δοκιμασία).

Η δοκιμασία F λαμβάνει υπόψη το λόγο των δύο μεταβλητοτήτων (ή διακυμάνσεων)

$$\frac{s_1^2}{s_2^2}$$

Ο λόγος αυτός είναι πάντοτε ≥ 1 . Αν η μηδενική υπόθεση H_0 ισχύει (δηλ. οι επαναληψιμότητες είναι ίδιες), τότε ο λόγος των δύο μεταβλητοτήτων είναι ~ 1 . Οι βαθμοί ελευθερίας για κάθε μεταβλητότητα είναι $n-1$.

Παράδειγμα: Δύο μέθοδοι συγκρίνονται μεταξύ τους για τον προσδιορισμό διαλυτού οξυγόνου (οκτώ προσδιορισμοί για κάθε μέθοδο):

Πρότυπη μέθοδος: 72 mg/l οξυγόνο και τυπική απόκλιση $s = 3,31$

Προτεινόμενη μέθοδος: 72 mg/l " και " " $s = 1,51$

Είναι η επαναληψιμότητα της προτεινόμενης μεθόδου σημαντικά μεγαλύτερη (δηλ. έχει μικρότερη μεταβλητότητα) από την αντίστοιχη της πρότυπης μεθόδου?

$$F = 3,31^2 / 1,51^2 = 4,8$$

Ενδιαφέρει αν η προτεινόμενη μέθοδος είναι πιο επαναλήψιμη από την πρότυπη μέθοδο. Εφαρμόζεται μονόπλευρη δοκιμασία.

Κρίσιμη τιμή $F_{7,7} = 3,787$ ($P = 0,05$). Επειδή $F_{7,7} < 4,8$, η μηδενική υπόθεση απορρίπτεται. Δηλαδή, η μεταβλητότητα της πρότυπης μεθόδου είναι σημαντικά μεγαλύτερη από αυτή της προτεινόμενης μεθόδου.

Άρα, η προτεινόμενη μέθοδος είναι πιο επαναλήψιμη (δηλ. έχει μικρότερη μεταβλητότητα) από την πρότυπη.

Πίνακας Α.3 Κρίσιμες τιμές F για μονόπλευρη δοκιμασία ($P = 0,05$)

v_2	v_1												
	1	2	3	4	5	6	7	8	9	10	12	15	20
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45
3	10,13	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812	8,786	8,745	8,703	8,660
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,912	5,858	5,803
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,678	4,619	4,558
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099	4,060	4,000	3,938	3,874
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677	3,637	3,575	3,511	3,445
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,284	3,218	3,150
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	3,073	3,006	2,936
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,913	2,845	2,774
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,788	2,719	2,646
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,687	2,617	2,544
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714	2,671	2,604	2,533	2,459
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646	2,602	2,534	2,463	2,388
15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,588	2,544	2,475	2,403	2,328
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538	2,494	2,425	2,352	2,276
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494	2,450	2,381	2,308	2,230
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456	2,412	2,342	2,269	2,191
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423	2,378	2,308	2,234	2,155
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393	2,348	2,278	2,203	2,124

v_1 = βαθμοί ελευθερίας για τον αριθμητή, v_2 = βαθμοί ελευθερίας για τον παρονομαστή.

Στο παράδειγμα για τον προσδιορισμό χρωμίου σε φυτικό ιστό υποθέσαμε ότι οι μεταβλητότητες των δύο μεθόδων δεν διαφέρουν σημαντικά.

Υπενθυμίζω:

Μέθοδος 1: μέση τιμή 1,48 ng/g τυπική απόκλιση 0,28 ng/g

Μέθοδος 2: μέση τιμή 2,33 ng/g τυπική απόκλιση 0,31 ng/g

$$\text{Άρα } F = 0,31^2 / 0,28^2 = 1,23$$

Δεν έχουμε σοβαρό λόγο να περιμένουμε ότι κάποια από τις δύο παρουσιάζει μεγαλύτερη μεταβλητότητα από την άλλη.

Άρα, η αμφίπλευρη δοκιμασία είναι πιο κατάλληλη.

$F_{4,4} = 9,605$ ($P = 0,05$) $> 1,23$. Συνεπώς, δεν υπάρχει σημαντική διαφορά μεταξύ των δύο μεταβλητοτήτων σε επίπεδο αξιοπιστίας 95%.

Μερικές από τις στατιστικές μεθόδους, όπως τα t-tests και η analysis of variance, προϋποθέτουν ομοιογένεια μεταβλητοτήτων.

Για να εκτιμηθεί η ισότητα των μεταβλητοτήτων για δύο ή περισσότερες σειρές δεδομένων μιας μεταβλητής χρησιμοποιείται, κατά περίπτωση, η **δοκιμασία Bartlett's** ή η **δοκιμασία Levene**.