ΣΤΑΤΙΣΤΙΚΗ

Στατιστική λέγεται ο κλάδος των Μαθηματικών ο οποίος <u>συγκεντρώνει στοιχεία</u> που αναφέρονται σε ένα σύνολο αντικειμένων, τα <u>ταξινομεί</u>, και τα <u>παρουσιάζει</u> σε κατάλληλη μορφή ώστε να μπορούν να αναλυθούν και να ερμηνευθούν .

Πληθυσμός μιας έρευνας λέγεται το σύνολο των αντικειμένων που εξετάζουμε ως προς ένα ή περισσότερα χαρακτηριστικά.

Κάθε στοιχείο του πληθυσμού ονομάζεται **άτομο ή μονάδα** του πληθυσμού.

Το πλήθος των ατόμων ενός πληθυσμού , ονομάζεται **μέγεθος** του πληθυσμού .

Ένα χαρακτηριστικό, ως προς το οποίο εξετάζεται ένας πληθυσμός, ονομάζεται **μεταβλητή**.

Οι μεταβλητές μπορεί να είναι :

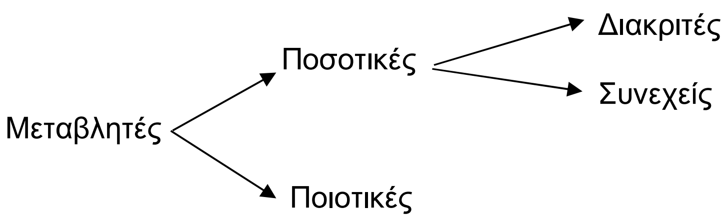
- **ποσοτικές,** τα οποία μπορούν να μετρηθούν (π.χ. αριθμός παιδιών μιας οικογένειας, ύψος ή βάρος ενός ατόμου, ηλικία κλπ)

ή

- **ποιοτικές (ή κατηγορικές)**, τα οποία δεν μπορούν να μετρηθούν αλλά περιγράφουν τα άτομα του πληθυσμού, π.χ. χρώμα ματιών, φύλο, κατάσταση υγείας κλπ.

Οι ποσοτικές μεταβλητές ειδικότερα διακρίνονται σε:
Διακριτές μεταβλητές, στις οποίες κάθε άτομο μπορεί να πάρει μόνο μεμονωμένες τιμές . (π.χ. αριθμός παιδιών, αριθμός βιβλίων που διάβασε κάποιος σε ένα διάστημα κλπ)

Συνεχείς μεταβλητές, στις οποίες κάθε άτομο μπορεί να πάρει οποιαδήποτε πραγματική τιμή που ανήκει σε κάποιο διάστημα ή ένωση διαστημάτων των πραγματικών αριθμών. (π.χ. βάρος, ύψος, ημερομίσθιο κλπ)
Σχηματικά έχουμε :



Απογραφή πληθυσμού λέγεται η καταγραφή και η μελέτη <u>όλων</u> των ατόμων του πληθυσμού ως προς μια ή περισσότερες μεταβλητές.
Επειδή η απογραφή είναι πολύ δύσκολη για τεχνικούς λόγους αλλά και δαπανηρή, επιλέγουμε να εξετάσουμε ένα μέρος του πληθυσμού το οποίο όμως πρέπει να είναι <u>αντιπροσωπευτικό</u> του πληθυσμού, δηλαδή να αποτελεί μικρογραφία του. Το μέρος αυτό του πληθυσμού ονομάζεται **δείγμα**, και η διαδικασία αυτή **δειγματοληψία**.
Αν το δείγμα δεν επιλεγεί σωστά, δε θα είναι αξιόπιστο και η έρευνά μας μπορεί να μην οδηγήσει σε σωστά αποτελέσματα
Ο αριθμός των ατόμων του δείγματος λέγεται **μέγεθος του δείγματος** και συμβολίζεται με ν.

ΠΑΡΟΥΣΙΑΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΜΕ ΠΙΝΑΚΕΣ

Έστω $x_1, x_2, \ldots, x_\kappa$ οι διαφορετικές τιμές μιας μεταβλητής X που αφορά ένα δείγμα μεγέθους ν, με $\kappa \leq \nu$.

- Αν όλες οι τιμές εμφανίζονται μία φορά, τότε κ = ν .
- Αν κ < ν , προφανώς κάποιες ή όλες οι τιμές εμφανίζονται περισσότερες από μία φορά.

⇨ Ο φυσικός αριθμός ν_i που δείχνει πόσες φορές εμφανίζεται η τιμή x_i της μεταβλητής X στο σύνολο των παρατηρήσεων λέγεται **συχνότητα** της παρατήρησης x_i .

Το άθροισμα των συχνοτήτων ν_i είναι ίσο με το μέγεθος του δείγματος ν

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

⇨ Το πηλίκο της συχνότητας μιας παρατήρησης προς το μέγεθος του δείγματος λέγεται **σχετική συχνότητα** της παρατήρησης x_i και συμβολίζεται με f_i .

Δηλαδή $\boxed{f_i = \dfrac{\nu_i}{\nu}}$ i = 1, 2, 3, …, κ

Για τη σχετική συχνότητα f_i της τιμής x_i ισχύει

1. $\boxed{0 \leq f_i \leq 1}$ για i = 1, 2, 3, …, κ
2. $\boxed{f_1 + f_2 + \ldots + f_\kappa = 1}$

Η σχετική συχνότητα f_i της τιμής x_i εκφράζει <u>τι μέρος του συνόλου των παρατηρήσεων αποτελούν οι παρατηρήσεις που είναι ίσες με x_i.</u>

Γι' αυτό, είναι χρήσιμο να την εκφράζουμε σε ποσοστό επί τοις εκατό και ισχύει

$$f_i\% = 100 f_i$$

και

$$f_1\% + f_2\% + \ldots + f_\kappa\% = 100$$

Για τις ποσοτικές μεταβλητές ορίζονται επίσης οι έννοιες :

Αθροιστική συχνότητα N_i της τιμής x_i , λέγεται το πλήθος των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i .

Είναι : $N_1 = \nu_1$

και $N_i = N_{i-1} + \nu_i$ ή $\boxed{N_i = \nu_1 + \nu_2 + \ldots + \nu_i}$

Αντίστοιχα έχουμε

Σχετική αθροιστική συχνότητα F_i της τιμής x_i, λέγεται το ποσοστό των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i .

Ισχύει : $F_1 = f_1$

και $F_i = F_{i-1} + f_i$ ή $F_i = f_1 + f_2 + \ldots + f_i$

Στις ποιοτικές μεταβλητές <u>δεν</u> έχουμε αθροιστικές συχνότητες.

- Ο πίνακας στον οποίο συγκεντρώνονται όλα τα στοιχεία μιας κατανομής, λέγεται **πίνακας συχνοτήτων** της κατανομής.
- Το σύνολο των ζευγών (x_i, v_i) αποτελεί την **κατανομή συχνοτήτων** της μεταβλητής.
- Το σύνολο των ζευγών (x_i, f_i) ή $(x_i, f_i\%)$ αποτελεί την **κατανομή σχετικών συχνοτήτων** της μεταβλητής.

ΟΜΑΔΟΠΟΙΗΣΗ ΠΑΡΑΤΗΡΗΣΕΩΝ

Όταν έχουμε συνεχή μεταβλητή ή διακριτή αλλά με πολύ μεγάλο πλήθος παρατηρήσεων, τότε ομαδοποιούμε τις παρατηρήσεις σε κλάσεις ίσου πλάτους.

- **Εύρος R** των παρατηρήσεων λέγεται η διαφορά της μικρότερης από τη μεγαλύτερη παρατήρηση.

- **αριθμός k των κλάσεων** εξαρτάται από το μέγεθος του δείγματος

Μέγεθος δείγματος ν	Αριθμός κλάσεων κ	Μέγεθος δείγματος ν	Αριθμός κλάσεων κ
<20	5	200-400	9
20-50	6	400-700	10
50-100	7	700-1000	11
100-200	8	≥1000	12

- Το πλάτος των κλάσεων c είναι $\boxed{c = \frac{R}{k}}$ στρογγυλοποιημένο αν χρειάζεται στον αμέσως μεγαλύτερο ακέραιο.
- Η κεντρική τιμή ή κέντρο μιας κλάσης [α,β) είναι $\boxed{x_i = \frac{\alpha + \beta}{2}}$
- Για κάθε κλάση [α,β) ισχύει $\boxed{c = \beta - \alpha}$
- Τα κέντρα δυο διαδοχικών κλάσεων διαφέρουν κατά το πλάτος c των κλάσεων.
- Καμία παρατήρηση δε θα μείνει έξω από κάποια κλάση
- Μια τιμή που συμπίπτει με το άνω όριο μιας κλάσης, μπαίνει στην επόμενη.

ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΚΑΤΑΝΟΜΗΣ ΣΥΧΝΟΤΗΤΩΝ

Α. Ποιοτικές Μεταβλητές

Ραβδογράμματα – Κυκλικό διάγραμμα – Σημειόγραμμα

- Το **ραβδόγραμμα** συχνοτήτων αποτελείται από ορθογώνιες στήλες (ράβδους), με βάση είτε στον οριζόντιο είτε στον κάθετο άξονα, που καθεμία αντιστοιχεί σε μία τιμή της μεταβλητής Χ και έχει ύψος ίσο με την αντίστοιχη συχνότητα.
  Το ραβδόγραμμα και μπορεί να αφορά συχνότητες ν_i, σχετικές συχνότητες f_i, ή σχετικές συχνότητες f_i%.

- Σε ένα **κυκλικό διάγραμμα**, στην παρατήρηση x_i αντιστοιχεί κυκλικός τομέας με μέτρο

$$a_i = f_i \cdot 360°$$

Επομένως ισχύει $a_i = \frac{\nu_i}{\nu} \cdot 360°$ άρα και $\frac{\nu_i}{\nu} = \frac{a_i}{360°}$

Οι τύποι αυτοί ισχύουν για το κυκλικό διάγραμμα σε όλα τα είδη των μεταβλητών

Β. Ποσοτικές Διακριτές Μεταβλητές

Διαγράμματα – Κυκλικό διάγραμμα - εικονογράμματα

- Το **διάγραμμα** συχνοτήτων μιας μεταβλητής Χ που οι τιμές της έχουν διαταχθεί σε αύξουσα σειρά στον οριζόντιο άξονα, αποτελείται από κατακόρυφες γραμμές που καθεμιά αντιστοιχεί σε μια τιμή x_i και έχει ύψος ίσο με την αντίστοιχη συχνότητα ν_i.

- Αν ενώσουμε τις κορυφές των γραμμών, δηλαδή τα σημεία (x_i, ν_i) προκύπτει το **πολύγωνο συχνοτήτων.**

  Το διάγραμμα και το πολύγωνο μπορεί να αφορούν συχνότητες ν_i, αθροιστικές συχνότητες N_i ,σχετικές συχνότητες f_i ή f_i%,
  ή και σχετικές αθροιστικές συχνότητες F_i ή F_i%.

Γ. Ομαδοποιημένες Κατανομές

Ιστογράμματα – Κυκλικό διάγραμμα

Για να κάνουμε το **ιστόγραμμα συχνοτήτων**, στον οριζόντιο άξονα του συστήματος των αξόνων τοποθετούμε με κατάλληλη κλίμακα τα όρια των κλάσεων. Στη συνέχεια κατασκευάζουμε διαδοχικά ορθογώνια, καθένα από τα οποία έχει βάση ίση με το πλάτος της κλάσης και ύψος ίσο με την αντίστοιχη συχνότητα.
Θεωρώντας το πλάτος c ως μονάδα στον οριζόντιο άξονα, το εμβαδό κάθε ορθογωνίου είναι ίσο με τη συχνότητα της κλάσης.

Πολύγωνο συχνοτήτων λέγεται το πολύγωνο που προκύπτει αν ενώσουμε τα <u>μέσα</u> των άνω βάσεων των ορθογωνίων του ιστογράμματος συχνοτήτων θεωρώντας και δυο υποθετικές κλάσεις στην αρχή και το τέλος με μηδενική συχνότητα (μηδενικό ύψος).

Το εμβαδό που ορίζεται από το πολύγωνο συχνοτήτων και τον οριζόντιο άξονα είναι ίσο με ν.

Όμοια ορίζεται το **πολύγωνο σχετικών συχνοτήτων** στο ιστόγραμμα σχετικών συχνοτήτων.

Το εμβαδό που ορίζεται από το πολύγωνο σχετικών συχνοτήτων και τον οριζόντιο άξονα είναι ίσο με 1.

Το εμβαδό που ορίζεται από το πολύγωνο σχετικών συχνοτήτων f_i% και τον οριζόντιο άξονα είναι ίσο με 100.

Το ιστόγραμμα και το πολύγωνο μπορεί να αφορούν συχνότητες v_i, αθροιστικές συχνότητες N_i ,σχετικές συχνότητες f_i ή f_i%, ή και σχετικές αθροιστικές συχνότητες F_i ή F_i%.

Πολύγωνο αθροιστικών συχνοτήτων λέγεται το πολύγωνο που προκύπτει αν ενώσουμε τα <u>δεξιά άκρα</u> των άνω βάσεων των ορθογωνίων του ιστογράμματος συχνοτήτων αρχίζοντας από το αριστερό άκρο της κάτω βάσης του πρώτου ορθογωνίου και τελειώνοντας στο δεξιό άκρο της άνω βάσης του τελευταίου .

Όμοια ορίζεται το **πολύγωνο αθροιστικών σχετικών συχνοτήτων** στο ιστόγραμμα αθροιστικών σχετικών συχνοτήτων.

<u>Χρονόγραμμα ή χρονολογικό διάγραμμα</u>

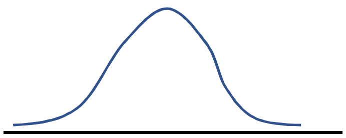
Χρησιμοποιείται για τη γραφική απεικόνιση της διαχρονικής εξέλιξης ενός οικονομικού , δημογραφικού ή άλλου μεγέθους.
Στον οριζόντιο άξονα ο χρόνος, στον κάθετο η μεταβλητή.

<u>Καμπύλες συχνοτήτων</u>

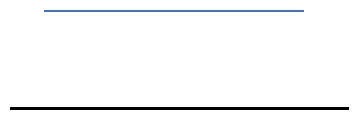
Αν ο αριθμός των κλάσεων σε μια συνεχή μεταβλητή γίνει πολύ μεγάλος, και το πλάτος των κλάσεων πολύ μικρό, τότε το πολύγωνο συχνοτήτων τείνει να πάρει τη μορφή μιας ομαλής καμπύλης, η οποία λέγεται καμπύλη συχνοτήτων.

Η μορφή της εξαρτάται από το πώς κατανέμονται οι παρατηρήσεις.

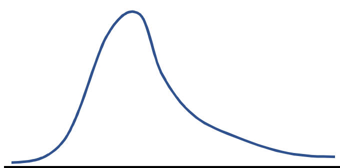
Χαρακτηριστικές μορφές είναι



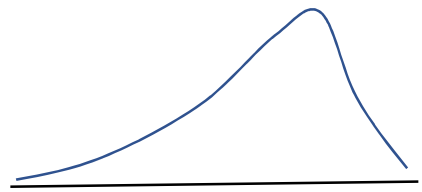
Κανονική κατανομή



Ομοιόμορφη κατανομή



Ασύμμετρη με θετική ασυμμετρία



Ασύμμετρη με αρνητική ασυμμετρία

ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ

Εκτός από τους πίνακες και τα διαγράμματα, υπάρχουν και κάποια αριθμητικά μέτρα με τα οποία μπορούμε να περιγράψουμε μια κατανομή.

Αυτά διακρίνονται σε

- **Μέτρα θέσης**, τα οποία δείχνουν τη θέση του κέντρου των παρατηρήσεων πάνω στον οριζόντιο άξονα

- **Μέτρα διασποράς (ή μεταβλητότητας)** που εκφράζουν τις αποκλίσεις των τιμών της μεταβλητής γύρω από τα μέτρα κεντρικής τάσης όπως είναι η μέση τιμή. (Πρακτικά το πόσο οι παρατηρήσεις εκτείνονται από το κέντρο.)

- **Μέτρα ασυμμετρίας** που καθορίζουν τη μορφή της καμπύλης συχνοτήτων και το αν αυτή παρουσιάζει συμμετρία . (δες τις παραπάνω καμπύλες συχνοτήτων)

ΜΕΤΡΑ ΘΕΣΗΣ → Μέση τιμή (ή αριθμητικός μέσος)
→ Σταθμικός μέσος (ή σταθμισμένος αριθμητικός μέσος)
→ Διάμεσος

Μέση τιμή $\overline{x}$

<u>Μέση τιμή</u> ενός συνόλου δεδομένων είναι το άθροισμα των παρατηρήσεων διά το πλήθος τους.

- Αν σε ένα δείγμα μεγέθους ν, οι παρατηρήσεις μιας μεταβλητής X είναι $t_1, t_2, \ldots, t_\nu$, (όλες διαφορετικές) τότε η μέση τιμή είναι:

$$\overline{x} = \frac{t_1 + t_2 + \ldots + t_\nu}{\nu} = \frac{\sum_{i=1}^{\nu} t_i}{\nu} \qquad \acute{\eta} \qquad \overline{x} = \frac{1}{\nu}\sum_{i=1}^{\nu} t_i$$

Στις ασκήσεις συνήθως χρησιμοποιούμε το συμβολισμό x_i και όχι t_i για τις παρατηρήσεις μας.

- Αν έχουμε κατανομή συχνοτήτων και η μεταβλητή X παίρνει τις διαφορετικές τιμές $x_1, x_2, \ldots x_\kappa$, με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots \nu_\kappa$ τότε η μέση τιμή είναι:

$$\overline{x} = \frac{x_1\nu_1 + x_2\nu_2 + \ldots + x_\kappa\nu_\kappa}{\nu_1 + \nu_2 + \ldots + \nu_\kappa} = \frac{\sum_{i=1}^{\kappa} x_i\nu_i}{\sum_{i=1}^{\kappa} \nu_i} = \frac{1}{\nu}\sum_{i=1}^{\kappa} x_i\nu_i$$

- Αν γνωρίζουμε τις σχετικές συχνότητες f_i, τότε

$$\overline{x} = \sum_{i=1}^{\kappa} x_i f_i$$

- Σε ομαδοποιημένες κατανομές η μέση τιμή δίνεται από τους ίδιους τύπους. Ως x_i παίρνουμε το κέντρο της i κλάσης.

  (Θυμίζουμε ότι το κέντρο μιας κλάσης [α, β) είναι $x_i = \frac{a+\beta}{2}$)

** Η μέση τιμή <u>επηρεάζεται από τις ακραίες τιμές</u> και <u>εξαρτάται από όλες τις τιμές</u> της μεταβλητής.

Σταθμικός μέσος $\overline{x}$

Αν οι ν παρατηρήσεις x_1, x_2,…, x_ν ενός συνόλου δεδομένων έχουν διαφορετικούς συντελεστές βαρύτητας (ή στάθμισης), τότε ο σταθμικός μέσος δίνεται από τον τύπο

$$\overline{x} = \frac{x_1 w_1 + x_2 w_2 + \ldots + x_\nu w_\nu}{w_1 + w_2 + \ldots + w_\nu} = \frac{\sum_{i=1}^{\nu} x_i w_i}{\sum_{i=1}^{\nu} w_i}$$

Διάμεσος (δ)

Η διάμεσος είναι η τιμή για την οποία το πολύ το 50% των παρατηρήσεων είναι μικρότερες από αυτήν και το πολύ το 50% των παρατηρήσεων είναι μεγαλύτερες από αυτήν.

Διάμεσος δ ενός δείγματος ν παρατηρήσεων που έχουν διαταχθεί σε αύξουσα σειρά, ονομάζεται:

- Η μεσαία παρατήρηση αν το πλήθος των παρατηρήσεων είναι περιττό. (μονός αριθμός)
- Ο μέσος όρος, δηλαδή το ημιάθροισμα, των δύο μεσαίων παρατηρήσεων αν το πλήθος των παρατηρήσεων είναι άρτιο. (ζυγός αριθμός)

** Η διάμεσος δεν επηρεάζεται από τις ακραίες τιμές .

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

- Εύρος (ή κύμανση) R
- Διακύμανση (ή διασπορά) s^2
- Τυπική απόκλιση s
- Συντελεστής Μεταβολής (ή Μεταβλητότητας)

Εύρος (R)

<u>**Εύρος R**</u> ενός συνόλου παρατηρήσεων ονομάζουμε τη διαφορά της μικρότερης τιμής από την μεγαλύτερη..

$$R = x_{\max} - x_{\min}$$

Το εύρος βασίζεται μόνο από στις δυο ακραίες παρατηρήσεις και γι' αυτό δε θεωρείται αξιόπιστο μέτρο διασποράς.

Διακύμανση (s^2)

<u>**Διακύμανση s^2**</u> ενός συνόλου παρατηρήσεων $t_1, t_2, \ldots t_\nu$, ονομάζουμε τη μέση τιμή των τετραγώνων των αποκλίσεων των t_i από τη μέση τιμή $\overline{x}$.

- Αν η μεταβλητή μας παίρνει τις ν διαφορετικές τιμές $t_1, t_2, \ldots t_\nu$ οι οποίες έχουν μέση τιμή $\overline{x}$, τότε

$$s^2 = \frac{1}{\nu}\sum_{i=1}^{\kappa}(t_i - \overline{x})^2$$

- Αν έχουμε πίνακα συχνοτήτων των παρατηρήσεων και οι ν παρατηρήσεις μας παίρνουν τις κ διαφορετικές τιμές $x_1, x_2, \ldots x_\kappa$ με αντίστοιχες συχνότητες $\nu_1, \nu_2, \ldots \nu_\kappa$, τότε η διακύμανση είναι

$$s^2 = \frac{1}{\nu}\sum_{i=1}^{\kappa}(x_i - \overline{x})^2 \nu_i$$

Προσοχή!
Στις ομαδοποιημένες κατανομές ισχύουν οι ίδιοι τύποι. Ως x_i λαμβάνονται τα μέσα x_i των κλάσεων.

*** Το μειονέκτημα της διακύμανσης είναι ότι οι μονάδες στις οποίες υπολογίζεται είναι το τετράγωνο της μονάδας της αντίστοιχης μεταβλητής. Γι' αυτό, αντί για τη διακύμανση , χρησιμοποιούμε ως μέτρο διασποράς την τετραγωνική της ρίζα

Τυπική απόκλιση (s)

<u>Τυπική απόκλιση s</u> ενός συνόλου παρατηρήσεων λέγεται η θετική τετραγωνική ρίζα της διακύμανσης των παρατηρήσεων , δηλ

$$s = \sqrt{s^2}$$

*** Η τυπική απόκλιση εκφράζεται στις ίδιες μονάδες με τη μεταβλητή!

Συντελεστής Μεταβολής CV (ή Μεταβλητότητας)

Όταν θέλουμε να συγκρίνουμε ως προς την ομοιογένεια δείγματα που είτε έχουν διαφορετική μέση τιμή και τυπική απόκλιση, είτε μετριούνται ως προς το ίδιο χαρακτηριστικό σε διαφορετικές μονάδες ή κλίμακες, τα μέτρα που έχουμε αναφέρει δεν επαρκούν. Για να ξεπεράσουμε αυτό το πρόβλημα ορίζουμε ένα ακόμη μέτρο διασποράς: το συντελεστή μεταβολής

Αν ένα δείγμα εξεταζόμενο ως προς μια ποσοτική μεταβλητή του, παρουσιάζει μέση τιμή $\bar{x} \neq 0$ και τυπική απόκλιση s , **<u>συντελεστής μεταβολής</u>** (ή συντελεστής μεταβλητότητας) ονομάζεται το πηλίκο:

$$CV = \frac{\tau\upsilon\pi\iota\kappa\eta\ \alpha\pi o\kappa\lambda\iota\sigma\eta}{\mu\varepsilon\sigma\eta\ \tau\iota\mu\eta} = \frac{s}{\bar{x}}$$

Αν η μέση τιμή είναι αρνητικός αριθμός τότε $\boxed{CV = \frac{s}{|\bar{x}|}}$

Ο συντελεστής μεταβολής

- είναι ένα μέτρο σχετικής διασποράς,
- δείχνει τη διασπορά απαλλαγμένη από την επίδραση της μέσης τιμής.
- είναι καθαρός αριθμός δηλαδή ανεξάρτητος από τις μονάδες μέτρησης.
- Ο συντελεστής μεταβολής μετρά την ομοιογένεια του δείγματος.

  Αν CV < 10%, τότε ο πληθυσμός του δείγματος θεωρείται **ομοιογενής**.

- Αν Α και Β δύο δείγματα με συντελεστές μεταβολής $\mathbf{CV_A}$ και $\mathbf{CV_B}$ τότε αν

  $\mathbf{CV_A < CV_B}$ το δείγμα Α παρουσιάζει μεγαλύτερη ομοιογένεια από το δείγμα Β

ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Αν οι παρατηρήσεις $x_1, x_2, \ldots x_\nu$, με μέση τιμή $\overline{x}$, διάμεσο δ και τυπική απόκλιση s, ακολουθούν **κατανομή κανονική ή περίπου κανονική**, τότε

- Το **68%** των παρατηρήσεων βρίσκεται στο διάστημα $(\overline{x}-s, \overline{x}+s)$
- Το **95%** των παρατηρήσεων βρίσκεται στο διάστημα $(\overline{x}-2s, \overline{x}+2s)$
- Το **99,7%** των παρατηρήσεων βρίσκεται στο διάστημα $(\overline{x}-3s, \overline{x}+3s)$
- Το εύρος R είναι ίσο με περίπου 6 τυπικές αποκλίσεις $R \approx 6s$
- $\delta = \overline{x}$