ΜΕΡΟΣ Α | Μέσος όρος, διακύμανση, τυπική απόκλιση

1. Βασικές έννοιες πειραματικής στατιστικής

1.1. Αριθμητικός μέσος όρος (Average)

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

όπου $x_1, x_2 \dots x_n$ οι τιμές του δείγματος
και n είναι οι τιμές του δείγματος

Παράδειγμα ο μέσος όρος του δείγματος
6, 7, 7, 9, 10, 11, 12, 14, 15, 17

$$\bar{x} = \frac{6+7+\dots 15+17}{10} = \frac{108}{10} = 10.8$$

1.2 Διακύμανση (Παραλλακτικότητα ή διασπορά)

Τύπος 1: $$s^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x}) + \dots + (x_n-\bar{x})^2}{n-1} = \frac{\Sigma(x_i-}{n-1}$$

όπου x_i είναι οι τιμές του δείγματος, $\bar{x}$ ο μέσος όρος και n ο αριθμός των μελών του δείγματος. Ο αριθμητής είναι το άθροισμα τετραγώνων (ΑΤ) των αποκλίσεων των τιμών του δείγματος από το μέσο τους όρο. Όσο πιο κοντά στο μέσο όρο θα είναι οι τιμές, τόσο πιο μικρές θα είναι οι αποκλίσεις τους από αυτόν και τόσο πιο μικρός θα είναι ο αριθμητής της, αντιστρόφως όσο πιο διασκορπισμένες γύρω από το μέσο όρο είναι οι τιμές του δείγματος τόσο μεγαλύτερος θα είναι ο αριθμητής.
Οι λογαριασμοί γίνονται πιο εύκολα με αυτόν τύπο

Τύπος 2: $$s^2 = \frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}$$

Ο παραπάνω τύπος είναι πιο εύχρηστος
Παράδειγμα:

Να υπολογισθεί η παρασλικη κόσμα των δειγμάτων: (A) 4, 7, 8, 10, 10, 12, 14, 15
(B) 1, 3, 5, 8, 10, 15, 17, 21

Μέσος όρος $\bar{X}_A = 10$
$\bar{X}_B = 10$

$$S_A^2 = \frac{(4^2+7^2+\ldots+15^2)-(4+7+\ldots+15)^2/8}{8-1} = \frac{894-(80)^2/8}{7} = 13{,}43$$

$$S_B^2 = \frac{(1^2+3^2+\ldots+21^2)-(1+3+\ldots 21)^2/8}{8-1} = \frac{1154-(80)^2/8}{7} = 50{,}57$$

Οι τιμές του πρώτου δειγματος (A) είναι συγκεντρωμένες γύρω απο τον μέσο όρο και έχει μικρότερη παραλλακτικότητα. Οι τιμες του δειγματος (B) είναι πιό διασκορπισμένες γύρω απο τον μέσο όρο.

Να βρεθεί η τυπική απόκλιση S (standard deviation)
Η τετραγωνική ρίζα της διακυμανσης S^2 ονομάζεται τυπική απόκλιση και
Ο τύπος είναι: Τυπική αποκλιση $S=\sqrt{S^2}$

ή $$S=\sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$$

Στο παραπάνω παράδειγμα

$S_A = 3{,}66$
$S_B = 7{,}11$

Οι όρα του τυπου υπολογισμού της παραλλακτικότητας εχουν αποκτήσει ιδιαίτερα ονόματα:
Αθροισμα τετραγώνων (AT) ονομάζεται η ποσότητα $\Sigma(x_i-\bar{x})^2$
Μέσο τετράγωνο (MT) ονομάζεται ή Διακύμανση S^2
Διορθωτικός όρος (ΔΟ) ονομάζεται η ποσότητα $(\Sigma x_i)^2/n$
Βαθμοί ελευθερίας (BE) ονομάζεται η ποσοτητα $n-1$

Συντελεστής Παραλλακτικότητας ή Παραλλακτικότητας
Coefficient of variation (cv)

Χρησιμοποιείται όταν θέλουμε να συγκρίνουμε την ακρίβεια ενός πειράματος ή παρόμοιων πειραμάτων που έγιναν στο παρελθόν.
Η ακρίβεια ενός πειράματος καθορίζεται από το μέγεθος του πειραματικού σφάλματος.
Επειδή η απευθείας σύγκριση των σφαλμάτων διαφορετικών πειραμάτων δεν είναι πάντοτε δυνατή γιατί τα πειράματα δεν έχουν το ίδιο μέγεθος πειραματικής μονάδας χρησιμοποιείται ο δείκτης $CV\%$ ο οποίος εκφράζει τις τιμές σε % με αποτέλεσμα να συγκρίνονται πειράματα σε διαφορετικές τοποθεσίες, χρονιές ή και με διαφορετικές μονάδες μέτρησης.

Τύπος: $$CV\% = \frac{\sqrt{S^2}}{\bar{x}} = \frac{S}{\bar{x}}$$

όπου S^2 η Διακύμανση και S η Τυπική απόκλιση
$\bar{x}$ ο Μέσος όρος

Παράδειγμα 1.
Έχουμε δύο δείγματα με μόνη την διαφορά ότι οι τιμές του πρώτου εκφράζονται σε mg ενώ του δεύτερου σε g.

(A)	mg	2000	3200	1800	4000	5000
(B)	g	2,0	3,2	1,8	4,0	5,0

Μέσοι όροι $\bar{x}_A = 3200$ Διακυμάνσεις $S_A^2 = 1.820.000$
$\bar{x}_B = 3,2$ $S_B^2 = 1,82$

$$CV(A) = \frac{\sqrt{1820.000}}{3200} \times 100 = 42,16\%$$
$$CV(B) = \frac{\sqrt{1,82}}{3,2} \times 100 = 42,16\%$$

Βλέπουμε ο συντελεστής CV% είναι ανεξάρτητος από την μονάδα μέτρησης

Παράδειγμα 1

Σε ένα πείραμα συγκρίθηκε η απόδοση 56 κλώνων μηδικής. Χρησιμοποιήθηκαν δύο είδων πειραματικά το ένα μια συνεχής γραμμή μήκους 3μ' το άλλο, ένα μεμονωμένο φυτό. Οι μέσοι όροι και οι διακυμάνσεις ήταν:

Μ.Ο. γραμμών $\bar{x} = 23{,}1$ Διακύμανση γραμμών $s^2 = 7{,}47$

Μ.Ο μεμονωμένων φυτών $\bar{x} = 3{,}17$ Διακύμανση $s^2 = 0{,}88$

Τα CV% είναι:

Γραμμές: $CV\% = \dfrac{\sqrt{s^2}}{\bar{x}} = \dfrac{\sqrt{7{,}47}}{23{,}1} \times 100 = 11{,}83\%$

Μεμονωμένα: $CV\% = \dfrac{\sqrt{0{,}88}}{3{,}17} \times 100 = 29{,}59\%$

Διαπιστώνουμε ότι το CV των μεμονωμένων φυτών είναι υψηλότερο δηλαδή τα μεμονωμένα φυτά έχουν υψηλότερη Διακύμανση (παραλλακτικότητα)

Τιμές CV% : 0–10% δηλώνουν υψηλή ακρίβεια
" : 10–20% " μέτρια "
" : 20–30% " μικρή "
" > 30% " πολύ μικρή "

Γενικά σε πειράματα αγρού δεν δεχόμαστε CV% > 33%

Ασκήσεις

1. Δίνεται το δείγμα

81, 85, 73, 56, 74, 62, 84, 86, 78, 93

Υπολογίστε για το δείγμα αυτό:

α) Τον μέσο όρο ($\bar{x}$)

β) Την διακύμανση (παραλλακτικότητα) (s^2)

γ) Την τυπική απόκλιση (S)
δ) Τον συντελεστή παραλλακτικότητας CV%