Τμήμα Επιστήμης Φυσικής Αγωγής & Αθλητισμού

Δημοκρίτειο Πανεπιστήμιο Θράκης

Στατιστική
στη Φυσική Αγωγή

(Ν161)

2η Διάλεξη

Κοκκότης Χρήστος, PhD

Στόχοι Διάλεξης

Κατανόηση βασικών εννοιών:

- Περιγραφική Στατιστική
- Μέσος όρος
- Διακύμανση
- Συχνότητες
- Τυπική απόκλιση
- Συντελεστής μεταβλητότητας

Τι είναι η Περιγραφική Στατιστική;

- Μέθοδος για ανάλυση και παρουσίαση δεδομένων.
- Βοηθά στην κατανόηση τάσεων και κατανομών.
- Μέτρα κεντρικής τάσης (μέση τιμή και διάμεσος)
- Χρήσιμη για την ανάλυση επιδόσεων, μετρήσεων αθλητών, μετρήσεις κατά την αποκατάσταση κ.α.

4

Παράμετροι θέσης

όταν θέλουμε να εκφράσουμε μια μεταβλητή με έναν αριθμό.

π.χ.

- αριθμητικός μέσος (μέση τιμή ή μέσος όρος)

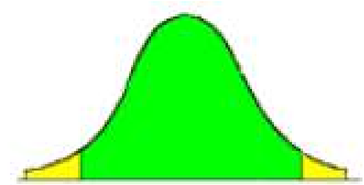
5

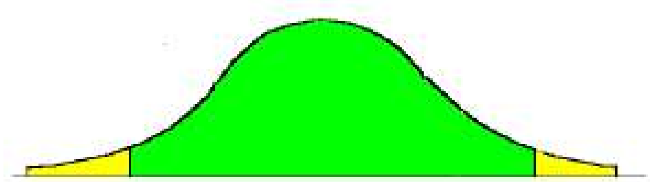
Παράμετροι διασποράς

όταν θέλουμε να εκφράσουμε τη συμπεριφορά μιας μεταβλητής με έναν αριθμό.

δηλ.

- πώς κατανέμεται,
- πού υπάρχει μεγαλύτερη συγκέντρωση τιμών,
- αν διαχέεται ομοιόμορφα και
- πόσο απλωμένες γύρω από τη μέση τιμή είναι οι τιμές της μεταβλητής





Παράμετροι διασποράς

όταν θέλουμε να εκφράσουμε τη συμπεριφορά μιας μεταβλητής με έναν αριθμό.

- **Διακύμανση**
- **Τυπική απόκλιση**
- **Συντελεστής μεταβλητότητας**
- **Εύρος μεταβολής**

Παράμετροι θέσης

Αριθμητικός μέσος

1. για μεμονωμένες τιμές (μιας μεταβλητής)
   - όχι ομαδοποιημένες τιμές
   - ούτε κατανομή συχνοτήτων

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_N}{N} = \frac{\sum_{i}^{N} x_i}{N}$$

X1, X2, X3, ..., XN= μεμονωμένες τιμές

N= αριθμός μεμονωμένων τιμών

π.χ. τιμές μεταβλητής **6,9,10,14, 16, 17**

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_N}{N} = \frac{\sum_{i}^{N} x_i}{N}$$

$$\bar{x} = \frac{\sum_{i}^{N} x_i}{N} = \frac{6+9+10+14+16+17}{6} = \frac{72}{6} = 12$$

Μέση τιμή των επιδόσεων 6 αθλητών σε ένα τεστ ευστοχίας

Παράμετροι θέσης

Αριθμητικός μέσος

2. για διακριτές μεταβλητές (κατανομή συχνοτήτων)
   - η κάθε τιμή της διακριτής μεταβλητής εμφανίζεται πολλές φορές

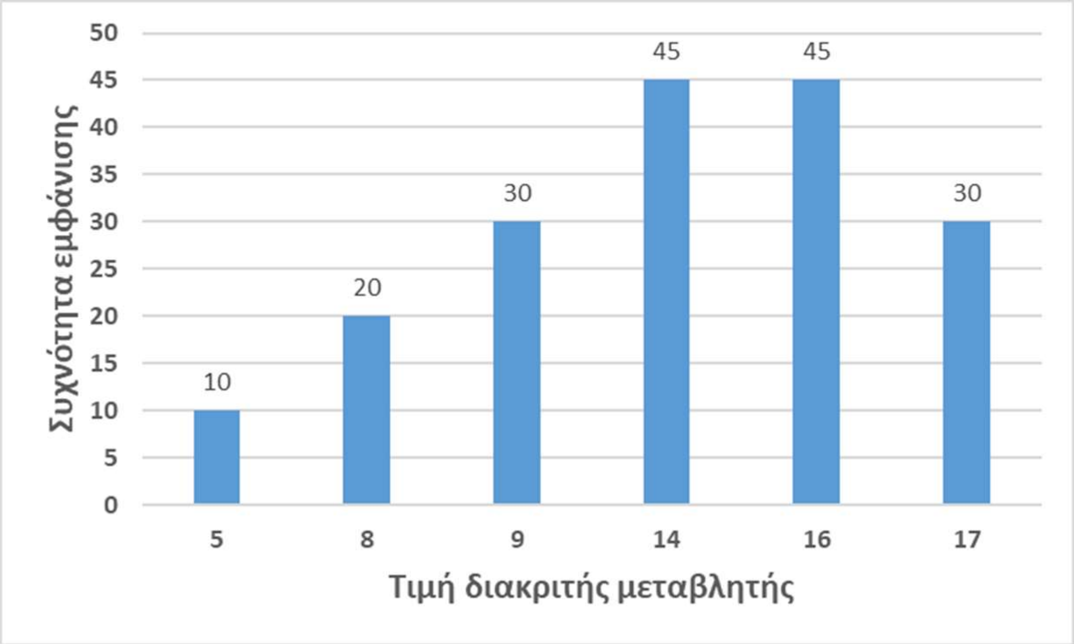
Τιμές Μεταβλητής	**Συχνότητα Εμφάνισης**
X_1	f_1
X_2	f_2
…	…
X_N	f_N

$$\overline{X} = \frac{\sum_{i}^{N} f_i \cdot x_i}{\sum_{i}^{N} f_i}$$

Άθροισμα συχνοτήτων εμφάνισης διακριτών τιμών

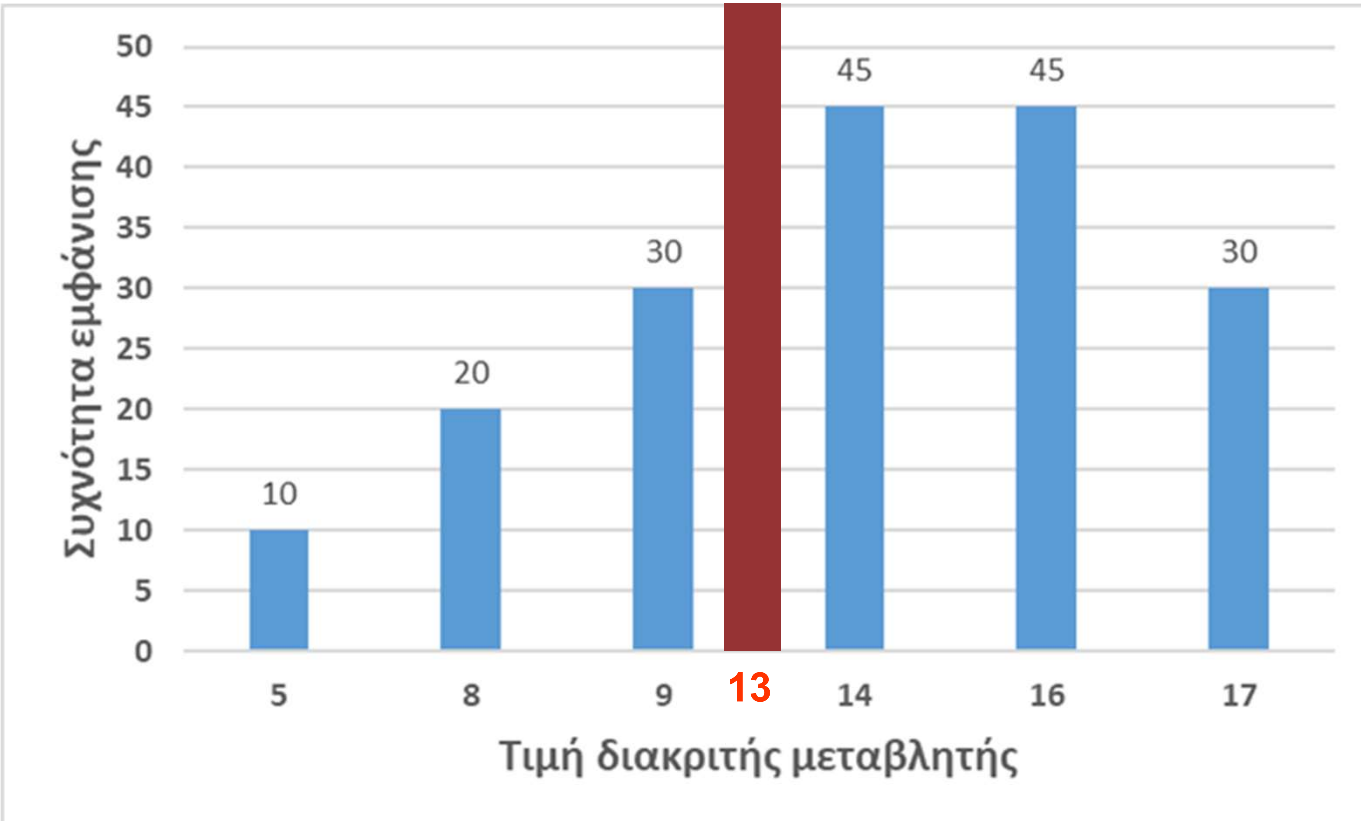
10

α/α	Τιμή διακριτής μεταβλητής (X_i)	Συχνότητα εμφάνισης (f_i)
1	5	10
2	8	20
3	9	30
4	14	45
5	16	45
6	17	30
N=6		



α/α	Τιμή διακριτής μεταβλητής (X_i)	Συχνότητα εμφάνισης (f_i)	$f_i \cdot \chi_i$
1	5	10	50
2	8	20	160
3	9	30	270
4	14	45	630
5	16	45	720
6	17	30	510
N=6		$\sum_{1}^{6} f_i = 180$	$\sum_{1}^{6} f_i \cdot \chi_i = 2340$

$$\overline{X} = \frac{\sum_{i}^{N} f_i \cdot x_i}{\sum_{i}^{N} f_i} \qquad \overline{X} = \frac{\sum_{i}^{N} f_i \cdot x_i}{\sum_{i}^{N} f_i} = \frac{2340}{180} = 13$$


50
45
40
35
30
25
20
15
10
5
0
Συχνότητα εμφάνισης
10
20
30
45
45
30
5
8
9
13
14
16
17
Τιμή διακριτής μεταβλητής

Παράμετροι θέσης

Αριθμητικός μέσος

3. για συνεχείς μεταβλητές

- τα δεδομένα εμφανίζονται με τη μορφή κατανομής συχνότητας κατά κλάσεις

κλάσεις	συχνότητα εμφάνισης
α_0-α_1	f_1
α_1-α_2	f_2
α_2-α_3	f_3
....	...

$$\overline{X} = \frac{\sum_i^k f_i \cdot x_i}{\sum_i^k f_i}$$

Άθροισμα συχνοτήτων εμφάνισης των κλάσεων

α_{k-1}-α_k= εύρος κλάσης

$X_i = \frac{a_i + a_{i-1}}{2}$ **= κεντρική τιμή της κλάσης α_{i-1}-α_i**

f_k= συχνότητα εμφάνισης της κλάσης

k= αριθμός κλάσεων

α/α	Κλάσεις	Συχνότητες f_i	X_i= Κεντρική τιμή κλάσης	$f_i \cdot X_i$
1	20-24	50	(24+20)/2=22	1100
2	25-27	60	(27+25)/2=26	1560
3	28-34	30	(34+28)/2=31	930
4	35-39	20	(35+39)/2=37	740
5	40-46	40	(46+40)/2=43	1720
K=5		$\sum_{1}^{5} f_i = 200$		$\sum_{1}^{5} f_i \cdot \chi_i = 6050$

$$\overline{X} = \frac{\sum_{i}^{k} f_i \cdot x_i}{\sum_{i}^{k} f_i} \qquad \overline{X} = \frac{\sum_{i}^{k} f_i \cdot x_i}{\sum_{i}^{k} f_i} = \frac{6050}{200} = 30.25$$

Παράμετροι διασποράς

- Διακύμανση
- Τυπική απόκλιση

πώς κατανέμεται μια μεταβλητή γύρω από το μέσο όρο

Απόκλιση της κάθε τιμής από το μέσο όρο;

α/α	X_i	X_i - $\bar{X}$
1	9	9-6= 3
2	8	8-6-= 2
3	8	8-6= 2
4	7	7-6= 1
5	5	5-6= -1
6	5	5-6= -1
7	4	4-6= -2
8	2	2-6= -4
	$\sum_{1}^{8} f_i = 48$	$\sum_{1}^{8} (X_i - \bar{X}) = 0$

$$\bar{X} = \frac{\sum_{i}^{N} X_i}{N} \qquad \bar{X} = \frac{\sum_{i}^{N} X_i}{N} = \frac{48}{8} = 6$$

Διακύμανση

Επειδή πάντα $\sum_{i}^{N}(X_i - \overline{X}) = 0$

γι' αυτό

υψώνουμε τις αποκλίσεις στο τετράγωνο και μετά αθροίζουμε

Διακύμανση= μέσος όρος των τετραγώνων των αποκλίσεων της κάθε τιμής (Xi) από το μέσο όρο ($\bar{X}$)

Παράμετροι διασποράς

Διακύμανση

1. για μεμονωμένες τιμές (μιας μεταβλητής)

- όχι ομαδοποιημένες τιμές
- ούτε κατανομή συχνοτήτων

Διακύμανση = $$\sigma^2 = \frac{\sum_{i}^{N}(X_i - \overline{X})^2}{N}$$

Διακύμανση

$$\overline{X} = \frac{\sum_{i}^{N} X_i}{N} = \frac{48}{8} = 6$$

α/α	$\mathbf{X_\iota}$	$(\mathbf{X_\iota} - \overline{\boldsymbol{X}})$	$(\mathbf{X_\iota} - \overline{\boldsymbol{X}})^2$
1	9	9-6= 3	$(3)^2$= 9
2	8	8-6= 2	$(2)^2$= 4
3	8	8-6= 2	$(2)^2$= 4
4	7	7-6= 1	$(1)^2$= 1
5	5	5-6= -1	$(-1)^2$= 1
6	5	5-6= -1	$(-1)^2$= 1
7	4	4-6= -2	$(-2)^2$= 4
8	2	2-6= -4	$(-4)^2$= 16
N= 8	$\sum_{1}^{8} X_i = 48$	$\sum_{1}^{8} (\mathrm{X}_i - \bar{\mathrm{X}}) = 0$	$\sum_{1}^{8} (\mathrm{X}_i - \bar{\mathrm{X}})^2 = 40$

Διακύμανση

Διακύμανση= $$\sigma^2 = \frac{\sum_{i}^{N}(X_i - \overline{X})^2}{N} = \frac{40}{8} = 5$$

Τυπική απόκλιση

ίδιες μονάδες μέτρησης όπως και οι παρατηρήσεις

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_i^N (X_i - \overline{X})^2}{N}}$$

Διακύμανση= $\sigma^2 = \frac{\sum_i^N (X_i - \overline{X})^2}{N} = \frac{40}{8} = 5$

Τυπική Απόκλιση= $\sigma = \sqrt{5}$ =2.23

Παράμετροι διασποράς

2. για διακριτές μεταβλητές (κατανομή συχνοτήτων)

- κάθε τιμή της διακριτής μεταβλητής εμφανίζεται πολλές φορές

Διακύμανση= $\sigma^2 = \dfrac{\sum_i^N f_i (X_i - \overline{X})^2}{\sum_i^N f_i}$

Άθροισμα συχνοτήτων εμφάνισης διακριτών τιμών

Τυπική απόκλιση= $\sigma = \sqrt{\dfrac{\sum_i^N f_i (X_i - \overline{X})^2}{\sum_i^N f_i}}$

Διακύμανση & Τυπική απόκλιση

$$\sigma^2 = \frac{\sum_{i}^{N} f_i (X_i - \overline{X})^2}{\sum_{i}^{N} f_i}$$

$$\overline{X} = \frac{\sum_{i}^{k} f_i \cdot x_i}{\sum_{i}^{k} f_i}$$

$$= \frac{2340}{180} = 13$$

$$\sum_{i}^{N} f_i (X_i - \overline{X}) = 0$$

α/α	X_i	f_i	$f_i \chi_i$	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$	$f_i \cdot (X_i - \overline{X})^2$
1	5	10	50	5-13= -8	64	640
2	8	20	160	8-13= -5	25	500
3	9	30	270	9-13= -4	16	480
4	14	45	630	14-13= 1	1	45
5	16	45	720	16-13= 3	9	405
6	17	30	510	17-13= 4	16	480
		180	2340			2550

Διακύμανση & Τυπική απόκλιση

$$\sigma^2 = \frac{\sum_i^N f_i (X_i - \overline{X})^2}{\sum_i^N f_i} = \frac{2550}{180} = 14.17$$

$$\sigma = \sqrt{\frac{\sum_i^N f_i (X_i - \overline{X})^2}{\sum_i^N f_i}} = \sqrt{14.17} = 3.76$$

Παράμετροι διασποράς

3. για συνεχείς μεταβλητές

- τα δεδομένα εμφανίζονται με τη μορφή κατανομής συχνότητας κατά κλάσεις

Διακύμανση= $$\sigma^2 = \frac{\sum_i^N f_i (X_{oi} - \overline{X})^2}{\sum_i^N f_i}$$

Τυπική απόκλιση= $$\sigma = \sqrt{\frac{\sum_i^N f_i (X_{oi} - \overline{X})^2}{\sum_i^N f_i}}$$

Διακύμανση & Τυπική απόκλιση

α/α	Κλάση	f_i	X_{oi}	$f_i \cdot \chi_{oi}$	$(X_{oi} - \bar{X})$	$(X_{oi} - \bar{X})^2$	$f_i \cdot (X_{oi} - \bar{X})^2$
1	20-24	50	22	1100	-8.25	68,0625	3403,125
2	25-27	60	26	1560	-4.25	18,0625	1083,75
3	28-34	30	31	930	0.75	0,5625	16,875
4	35-39	20	37	740	6.75	45,5625	911,25
5	40-46	40	43	1720	12.75	162,5625	6502,5
		200		6050			11917,5

$$\sum_{i}^{N} f_i (X_{oi} - \bar{X}) = 0 \qquad \sigma^2 = \frac{\sum_{i}^{N} f_i (X_{oi} - \bar{X})^2}{\sum_{i}^{N} f_i} \qquad \bar{X} = \frac{\sum_{i}^{k} f_i \cdot x_{oi}}{\sum_{i}^{k} f_i} = 30.25$$

Διακύμανση & Τυπική απόκλιση

Διακύμανση= $$\sigma^2 = \frac{\sum_i^N f_i (X_{oi} - \overline{X})^2}{\sum_i^N f_i} = \frac{11917.5}{200} = 59.59$$

Τυπική απόκλιση= $$\sigma = \sqrt{\frac{\sum_i^N f_i (X_{oi} - \overline{X})^2}{\sum_i^N f_i}} = \sqrt{59.59} = 7.72$$

Διακύμανση και τυπική απόκλιση δείγματος

1. για μεμονωμένες τιμές (μιας ποσοτικής μεταβλητής)

$$s^2 = \frac{\sum_{i}^{N} (X_i - \overline{X})^2}{N - 1} \qquad s = \sqrt{\frac{\sum_{i}^{N} (X_i - \overline{X})^2}{N - 1}}$$

2. για συχνότητες (μιας διακριτής μεταβλητής)

$$s^2 = \frac{\sum_{i}^{N} f_i (X_i - \overline{X})^2}{\sum_{i}^{N} f_i - 1} \qquad s = \sqrt{\frac{\sum_{i}^{N} f_i (X_i - \overline{X})^2}{\sum_{i}^{N} f_i - 1}}$$

Διακύμανση και τυπική απόκλιση δείγματος

3. για συχνότητες κατά κλάσεις (μιας συνεχής ποσοτικής μεταβλητής)

$$s^2 = \frac{\sum_{i}^{N} f_i (X_{oi} - \overline{X})^2}{\sum_{i}^{N} f_i - 1}$$

$$s = \sqrt{\frac{\sum_{i}^{N} f_i (X_{oi} - \overline{X})^2}{\sum_{i}^{N} f_i - 1}}$$

Συντελεστής μεταβλητότητας

- σύγκριση δύο κατανομών (πληθυσμών ή δειγμάτων) ως προς την ομοιογένειά τους
- ποιο είναι το πιο ομοιογενές δείγμα; (δηλ. το δείγμα με τη μικρότερη διασπορά)

$$CV(x) = \frac{\sigma}{\mu} \qquad CV(x) = \frac{s}{\overline{X}}$$

σ= τυπική απόκλιση πληθυσμού
μ= μέσος όρος πληθυσμού

s= τυπική απόκλιση δείγματος
X= μέσος όρος δείγματος

Συντελεστής μεταβλητότητας

$$CV(x) = \frac{s}{\overline{X}}$$

Μικρός συντελεστής μεταβλητότητας

⇩

Μικρή Διασπορά

⇩

Μεγάλη ομοιογένεια

Σύγκριση δύο δειγμάτων

1ο δείγμα:

$$\overline{X} = 120$$

$$s = 20$$

$$CV(x_1) = \frac{s}{\overline{X}} = \frac{20}{120} = 0.166$$

2ο δείγμα:

$$\overline{X} = 120$$

$$s = 10$$

$$CV(x_2) = \frac{s}{\overline{X}} = \frac{10}{120} = 0.083$$

το 2ο δείγμα είναι πιο ομοιογενές (0.083 < 0.166)

Σύγκριση δύο δειγμάτων

1ο δείγμα: $\overline{X} = 80$

$s = 20$

$$CV(x_1) = \frac{s}{\overline{X}} = \frac{20}{80} = 0.25$$

2ο δείγμα: $\overline{X} = 70$

$s = 10$

$$CV(x_2) = \frac{s}{\overline{X}} = \frac{10}{70} = 0.143$$

το 2ο δείγμα είναι πιο ομοιογενές (0.143 < 0.25)

Σύγκριση δύο δειγμάτων

1ο δείγμα: $\overline{X} = 80$

$s = 20$

$$CV(x_1) = \frac{s}{\overline{X}} = \frac{20}{80} = 0.25$$

2ο δείγμα: $\overline{X} = 70$

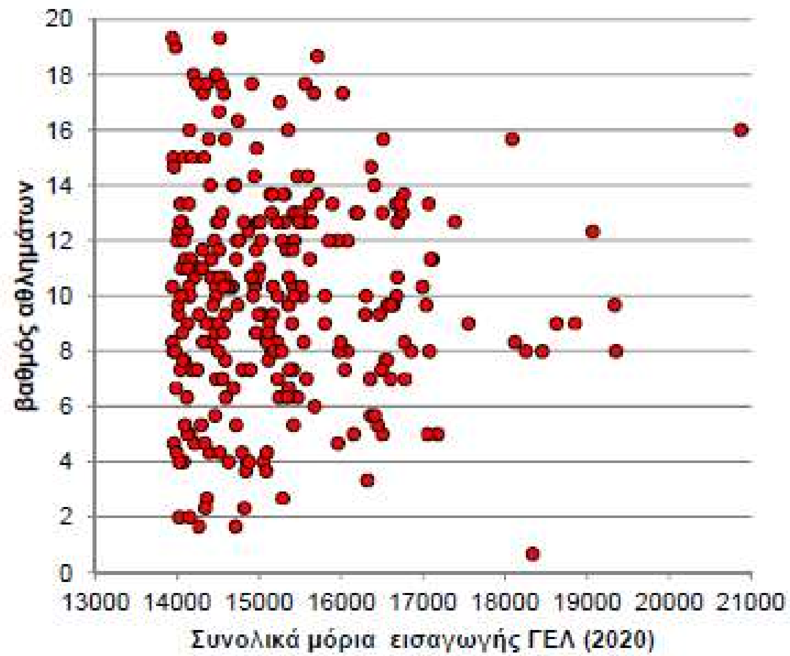
$s = 20$

$$CV(x_2) = \frac{s}{\overline{X}} = \frac{20}{70} = 0.286$$

το 1ο δείγμα είναι πιο ομοιογενές (0.25 < 0.286)

Εύρος Μεταβολής

min – max
0.67 – 19.33



min – max
13941 – 20875

Εύρος Μεταβολής (R) = Μεγαλύτερη παρατήρηση – Μικρότερη Παρατήρηση

Εφαρμογές στην Αθλητική Επιστήμη

- Ανάλυση επιδόσεων αθλητών.
- Αξιολόγηση της αποτελεσματικότητας προπονητικών μεθόδων.
- Μελέτη φυσιολογικών μεταβλητών (καρδιακός ρυθμός, VO2 max, γαλακτικό κατώφλι).

Σύνοψη

- Ο Μέσος Όρος χρησιμοποιείται για την εύρεση της κεντρικής τιμής.
- Η Διακύμανση και η Τυπική Απόκλιση δείχνουν τη μεταβλητότητα των δεδομένων.
- Ο Συντελεστής μεταβλητότητας ελέγχει την ομοιογένεια του δείγματος.

Ερωτήσεις & Συζήτηση