ALL BRANCHES

ME,CE,EC,EE,CS

Probability and Statistics

Lecture No- 03

By- Vinay sir

Topics to be Covered

01 Introduction to statistics

02 Ungrouped data

03 Grouped data

04 Introduction to Random variables

Revision

Probability → Random Experiments

$$\frac{n(A)}{n(S)}$$

$$P(A/B) \cdot P(B) = P(B/A) \cdot P(A)$$

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

Conditional Probability → $P(A/B)$

If $E_1, E_2, E_3, \ldots E_n$ are Such that $E_i \cap E_j = \phi$ and $E_1 \cup E_2 \cup E_3 \cup \ldots E_n = S$, then

$$P(A) = \sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)$$

$$P(E_i/A) = \frac{P(E_i) \cdot P(A|E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)}$$

Introduction to Statistics

Statistics → collection and analysis of Data.

Survey → Data (ungrouped Data (or) Raw Data).

ungrouped Data → Raw Data

is converted to Grouped Data.

Ex: The Marks of 40 Students are.

19	24	31	32	17	16	9	22	24	17
14	23	20	06	20	19	06	22	27	11
06	29	14	12	18	32	30	17	13	19
23	11	33	31	28	24	21	09	11	30

Ungrouped data

(i) **Mean:** If $x_1, x_2, x_3, x_4, \ldots x_n$ are 'n' observations, then the Mean is denoted by μ or $\bar{x}$ and it is given by

$$\mu = \bar{x} = \frac{\text{Sum of observations}}{\text{No. of observations}} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Ex: The Mean of 1st 'n' natural numbers is ___.

$$\bar{x} = \frac{1+2+3+\cdots+n}{n} = \frac{\cancel{n}(n+1)}{2\cdot\cancel{n}} = \frac{n+1}{2}.$$

② **Median:** If $x_1, x_2, x_3, \ldots, x_n$ are 'n' observations, arranged in ascending (or) Descending order, then

(i) Median is $\left(\frac{n+1}{2}\right)^{th}$ observation if 'n' is odd.

(ii) Median is mean of $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n}{2}+1\right)^{th}$ observation if 'n' is even

Ex: let the observations be 19, 23, 27, 11, 14, 18, 40, 26, 27

Ascending order → 11, 14, 18, 19, 23, 26, 28, 2023, 2048.

$\therefore$ Median $= \left(\frac{9+1}{2}\right)^{th}$ observation $= 5^{th}$ observation $= 23$.

→ The Median of the data → 11, 23, 17, 12, 14, 28, 27, 13, 12, 20.

Ascending order → 11, 12, 12, 13, 14, 17, 20, 23, 27, 28.

Median is average of $\left(\frac{10}{2}\right)^{th}$ & $\left(\frac{10}{2}+1\right)^{th}$ observations.

$$\Rightarrow \text{Median} = \frac{14+17}{2} = \frac{31}{2} = 15.5$$

→ Mode: The most repeatitive data point in a data is called mode.

Ex: 2, 3, 5, 2, 7, 2, 14, 18, 23, 2, 12 → Mode = 2.

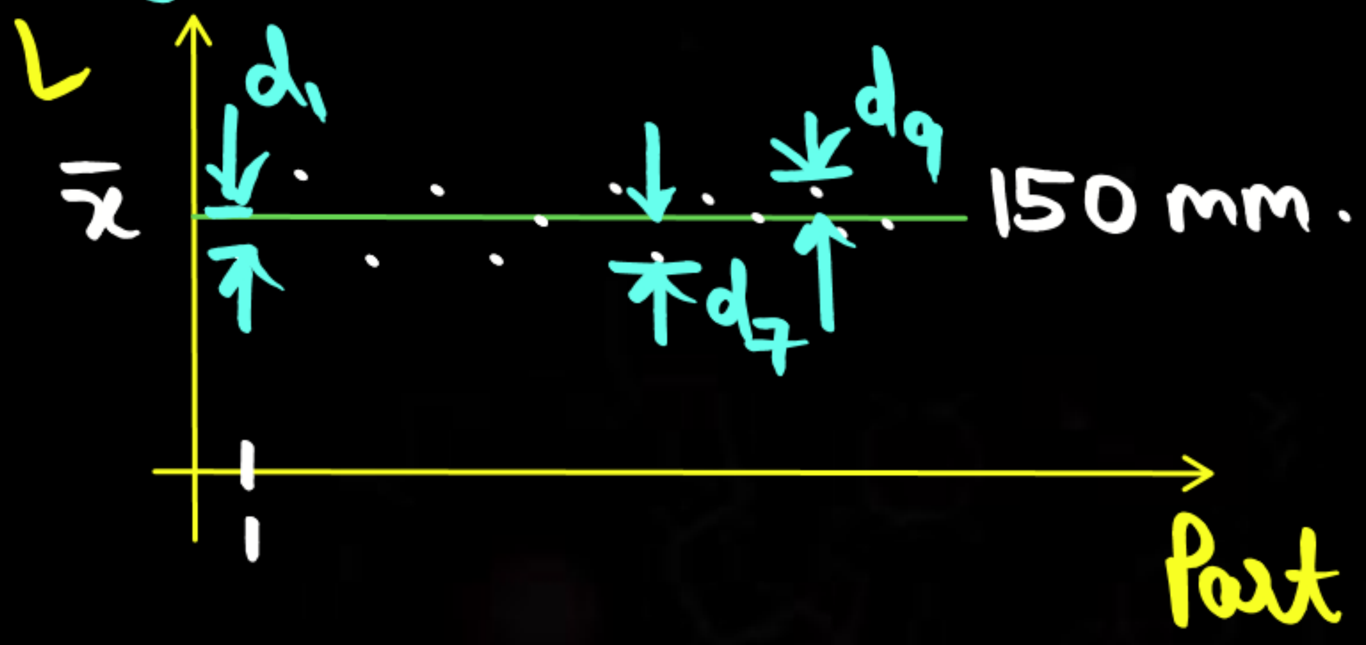
④ Deviation: (d_i):

let $x_1, x_2, x_3, x_4, \ldots, x_n$ are 'n' observations with mean $\bar{x}$, then $d_1, d_2, d_3, \ldots, d_n$ be the corresponding 'n' deviations.

Such that $d_1 = x_1 - \bar{\mu}$

$d_2 = x_2 - \bar{\mu}$

$\vdots$



Sum of deviations about Mean $= \sum_{i=1}^{n} (x_i - \bar{x}) = (x_1 + x_2 + x_3 + \cdots + x_n) - n \cdot \bar{x}$

$= \cancel{n \cdot \bar{x}} - \cancel{n \cdot \bar{x}} = 0.$

Variance: If $x_1, x_2, x_3, \ldots, x_n$ are 'n' data Points, then Variance of the data is denoted by σ^2 and it is given by

$$\sigma^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} d_i^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \mu)^2$$

Standard Deviation $(\sigma) = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \mu)^2}$

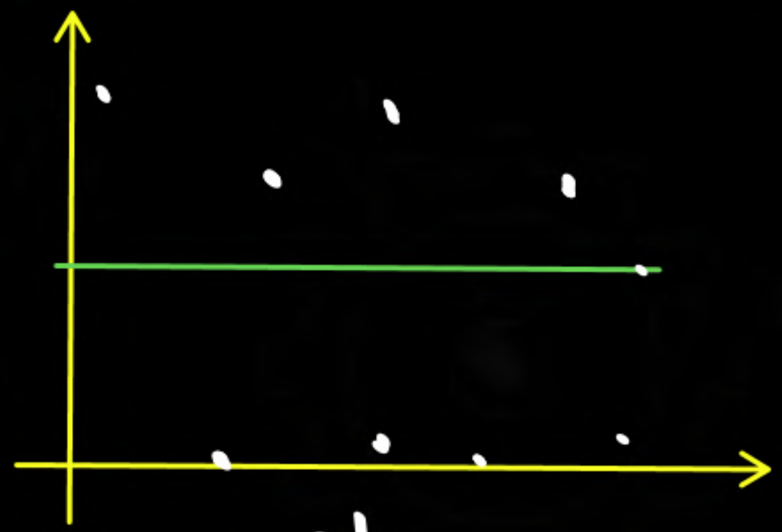
$$\boxed{\therefore \sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \mu)^2}} \longrightarrow \text{Standard Deviation}$$

Coefficient of Variation:

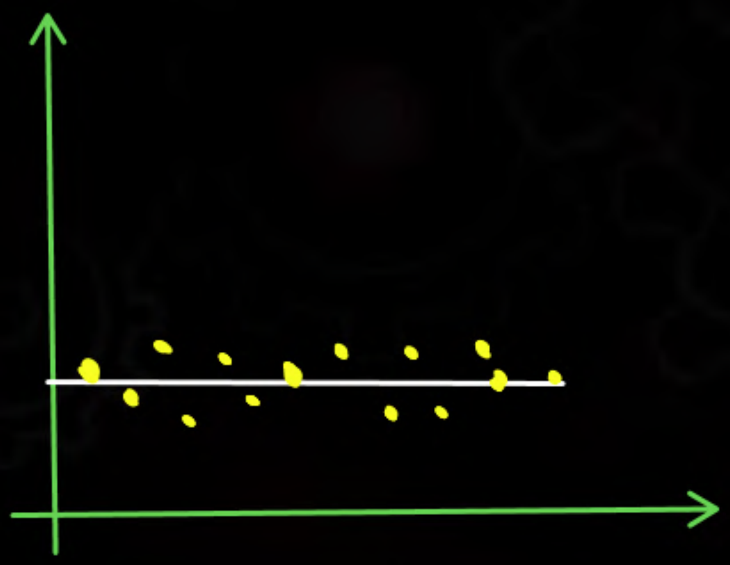
$$\text{Coefficient of Variation} = \frac{\sigma}{\mu}$$

Player-1 → High Average and high Standard deviation.

Player-2 → Less Average and also less Standard deviation



Player-1



Player-2

Ex: The Marks of 40 Students are.

19	24	31	32	17	16	9	22	24	17.
14	23	20	06	20	19	06	22	27	11
06	29	14	12	18	32	30	17	13	19
23	11	33	31	28	24	21	09	11	30

Grouped Data

unGrouped Data → Grouped Data.

Data is distributed into finite Number of classes.

class	frequency
0 - 5	0
5 - 10	5
10 - 15	7
15 - 20	8
20 - 25	10
25 - 30	3
30 - 35	7

Lower boundary

upper boundary

class width = upper boundary - Lower boundary.

Mean :

class	frequency (f_i)	class Mark (x_i)	$f_i x_i$
0-5	0	2.5	0
5-10	5	7.5	37.5
10-15	7	12.5	87.5
15-20	8	17.5	140
20-25	10	22.5	225
25-30	3	27.5	82.5
30-35	7	32.5	227.5

$\Sigma f = 40 = N$

$\Sigma f_i x_i = 800$

$$\therefore \bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

$$= \frac{800}{40} = 20$$

$$\therefore \bar{x} = 20$$

Median:

class	frequency	Cumulative frequency
0-5	0	0
5-10	5	5
10-15	7	12
15-20	8	20
20-25	10	30
25-30	3	33
30-35	7	40

$\Sigma f = 40 = N$

Median Class

$$\text{Median} = L + \frac{\left(\frac{N}{2} - F\right)}{f} \times C$$

$L \rightarrow$ Lower boundary of Median class

$N \rightarrow \Sigma f$;

$F \rightarrow$ Cumulative frequency of class preceeding the median class

$f \rightarrow$ frequency of Median Class

$C \rightarrow$ class width

$$\text{Median} = L + \frac{\left(\frac{N}{2} - F\right)}{f} \times C$$

$$= 15 + \left(\left(\frac{20-12}{8}\right) \times 5\right)$$

$$= 15 + \left(\frac{8}{8} \times 5\right) = 15 + 5 = \underline{20}$$

$$\boxed{\text{Median} = 20.}$$

Mode:

class	frequency
0-5	0
5-10	5
10-15	7
15-20	$f_1 = 8$
20-25	$f = 10$ → Modal class
25-30	$f_2 = 3$
30-35	7
	$\Sigma f = 40$

$$\text{Mode} = L + \frac{(f-f_1)}{(f-f_1)+(f-f_2)} \times C$$

L → Lower Boundary of Modal class.

f → frequency of Modal class.

f_1 → frequency of class Preceeding Modal class

f_2 → frequency of class Succeeding Modal Class.

C → Class width.

$$\text{Mode} = 20 + \frac{(10-8)}{(10-8)+(10-3)} \times 5$$

$$= 20 + \left(\frac{2}{2+7}\right) \times 5$$

$$= 20 + \left(\frac{10}{9}\right) = 21.11$$

$$\boxed{\text{Mode} = 21.11}$$

$$\left.\begin{array}{l} 10-14 \\ 15-19 \\ 20-24 \\ 25-29 \end{array}\right\} \rightarrow \text{Exclusive}$$

$$\downarrow$$

10–14.5
14.5–19.5
19.5–24.5
$24.5 \rightarrow 29$

Inclusive

$$\text{Median} = 3\,\text{Mean} - 2\,\text{Mode}.$$

→ Empirical formulae.

Thank You!

GW Soldiers