PROBABILITY AND STATISTICS (UCS401)

Lecture-1-2 Contd…
(Prerequisite of Statistics (Mean, S.D., C.V.))
Introduction to Statistics and Data Analysis

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Dr. Rajanish Rai

Assistant Professor

School of Mathematics

Thapar Institute of Engineering and Technology, Patiala

Pumpkin Sir

[7] Prerequisite of distribution theory (Mean, variance and standard deviation)

(1) Standard deviation for individual series –:

Standard deviation (σ) : (Root mean square deviation)

* $$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$$ (Actual mean method)

or $$\sqrt{\frac{\sum f(x_i-\bar{x})^2}{\sum f}}$$

where $$\bar{x} = \frac{\sum x}{n}$$

* $$\sigma = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2}$$ (Assumed mean method)

Here, $d = x - A$ & A = Assumed mean

and $$\boxed{\bar{x} = A + \frac{\sum d}{n}}$$

* Coefficient of standard deviation $= \dfrac{\sigma}{x}$

* Varience (σ^2) $= \dfrac{\sum (x-\bar{x})^2}{n}$

* $$\boxed{\text{Coefficient of variation} = \frac{\sigma}{x} \times 100}$$

Example - 1

Individual series,

10, 12, 13, 15, 20

$$\bar{x} = \frac{\sum x}{n} = \frac{10+12+13+15+20}{5} = \frac{70}{5} = 14$$

$$\boxed{\bar{x} = 14}$$

We know that by Actual mean method (Direct method)

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

Here, $\bar{x} = 14$ & $n = 5$

x	$x-\bar{x}$	$(x-\bar{x})^2$
10	-4	16
12	-2	4
13	-1	1
15	1	1
20	6	36
		$\sum (x-\bar{x})^2 = 58$

$$S.D. (\sigma) = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$$

$$= \sqrt{\frac{58}{5}} = \sqrt{11.6}$$

$$\boxed{\sigma = 3.4}$$

$\therefore$ Coefficient of standard deviation $= \frac{\sigma}{\bar{x}}$

$$= \frac{3.4}{14} = 0.2428.$$

$$\boxed{\text{Variance } (\sigma^2) = 11.6}$$

and Coefficient of variation $= \frac{\sigma}{\bar{x}} \times 100$

$$= 24.28\%$$

(b) Assumed mean method (short-cut method)

We know that

$$\sigma = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2} \qquad d = x - A$$

$A \downarrow$ Assumed mean.

Given. $n = 5$

x	$d = x - A$	$d^2 = (x-A)^2$
10	-3	9
12	-1	1
A ← 13	0	0
15	2	4
20	7	49
	$\sum d = 5$	$\sum d^2 = 63$

$$\sigma = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2} = \sqrt{\frac{63}{5} - \left(\frac{5}{5}\right)^2}$$

$$= \sqrt{12.6 - 1}$$

$$= \sqrt{11.6}$$

$$\boxed{\sigma = 3.4} \Rightarrow \text{same.}$$

Example (2) Find the standard deviation for individual series 48, 43, 65, 57, 31, 60, 37, 48, 59, 78.

$$\bar{x} = \frac{\sum x}{n} = \frac{48+43+65+57+31+60+37+48+59+78}{10}$$

$$= \frac{526}{10}$$

$= 52.6 \Rightarrow$ proceed through assumed mean method.

We know that

$$S.D.\,(\sigma) = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2} \; ; \; d = x - A$$

$A \downarrow$ assumed mean.

x	$d = x-A$	$d^2 = (x-A)^2$
31	-17	289
37	-11	121
43	-5	25
48	0	0
A ← 48	0	0
57	9	81
59	11	121
60	12	144
65	17	289
78	30	900
	$\sum d = 46$	$\sum d^2 = 1970$

$$\sigma = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2} = \sqrt{\frac{1970}{10} - \left(\frac{46}{10}\right)^2}$$

$$\sigma = \sqrt{175.84} = 13.26.$$

$$\boxed{S.D.\,(\sigma) = 13.26}$$

$$\text{Coefficient of } S.D. = \frac{\sigma}{\bar{x}} = \frac{13.26}{52.6} = 0.2526$$

$$\boxed{\text{Variance}(\sigma^2) = 175.84}$$

$$\text{Coefficient of variation} = \frac{\sigma}{\bar{x}} \times 100$$

$$= 25.26\%$$

(2) Standard deviation and variance for discreate series -:

(a) Assumed mean method -:

$$\sigma = \sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2}$$

Here, $d = x - A$, $\bar{x} = A + \frac{\sum fd}{\sum f}$

$\downarrow$ Assumed mean

(b) step-deviation method -:

$$\sigma = \left(\sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2}\right) \times h$$

Here, $d = \frac{x-A}{h}$ $\quad h \rightarrow$ Class interval (diff same)

$$\bar{x} = A + \left(\frac{\sum fd}{\sum f}\right) \times h$$

$A = 5$ (say)

Question -:

x	f	$d = x - A$	$\cancel{f}d^2$	fd	fd^2
1	1	-4	16	-4	16
3	2	-2	4	-4	8
5	3	0	0	0	0
7	4	2	4	8	16
8	5	3	9	15	45
	$\sum f = 15$			$\sum fd = 15$	$\sum fd^2 = 85$

We know that

$$S.D.(\sigma) = \sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2}$$

$$= \sqrt{\frac{85}{15} - \left(\frac{15}{15}\right)} = \sqrt{5.66 - 1} = \sqrt{4.66}$$

$$\boxed{S.D.(\sigma) = 2.16}$$

$$\bar{x} = A + \frac{\Sigma fd}{\Sigma f} = 5 + \frac{15}{15} = 6$$

$$\boxed{\bar{x} = 6}$$

$\therefore$ Coefficient of S.D. $= \frac{\sigma}{\bar{x}} = \frac{2.16}{6} = 0.36$.

$$\boxed{\text{Variance } (\sigma^2) = 4.66}$$

Coefficient of variation (CV) $= \frac{\sigma}{\bar{x}} \times 100 = 36\%$ Ans

Question-: The annual salaries of a group of employees are given in the following Table. Find S.D. & CV.

Salaries (in thous)	No. of persons
45	3
50	5
55	8
60	7
65	9
70	7
75	4
80	7

$\Rightarrow$ fixed gap $h = 5$.

we know that by step-deviation method

$$\sigma = \sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2} \times h$$

Here $d = \frac{x-A}{h}$ $h \to$ class interval.

$$\bar{x} = A + \left(\frac{\sum fd}{\sum f}\right) \times h$$

Here $h = 5$

x	f	$d = \frac{x-60}{5}$	d^2	fd	fd^2
45	3	-3	9	-9	27
50	5	-2	4	-10	20
55	8	-1	1	-8	8
A → 60	7	0	0	0	0
65	9	1	1	9	9
70	7	2	4	14	28
75	4	3	9	12	36
80	7	4	16	28	112
	$\sum f = 50$			$\sum fd = 36$	$\sum fd^2 = 240$

$$S.D.(\sigma) = \sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2} \times h$$

$$= \sqrt{\frac{240}{50} - \left(\frac{36}{50}\right)^2} \times 5$$

$$\boxed{\sigma = 10.35}$$

$$\bar{x} = A + \left(\frac{\sum fd}{\sum f}\right) \times h = 60 + \left(\frac{36}{50}\right) \times 5 = 60 + 3.6$$

$$\boxed{\bar{x} = 63.6}$$

Now Coefficient of variation (CV) $= \frac{\sigma}{\bar{x}} \times 100$

$$= \frac{10.35}{63.6} \times 100$$

$$\boxed{CV = 16.27\%}$$

Variance $(\sigma^2) = (10.35)^2 = 107.12$ Ans

(iii) For Continuous series -:

By step- deviation method,

$$\sigma = \left(\sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2}\right) \times h$$

Here, $d = \frac{x-A}{h}$, $h \rightarrow$ class interval

$$\bar{x} = A + \left(\frac{\Sigma fd}{\Sigma f}\right) \times h$$

Question -: Find S.D. and CV

Class (x)	f
0-10	15
10-20	15
20-30	23
30-40	22
40-50	25
50-60	10
60-70	5
70-80	10

Class interval (x)	x (mid)	f	$d=\frac{x-35}{10}$	d^2	fd	fd^2
0–10	5	15	-3	9	-45	135
10–20	15	15	-2	4	-30	60
20–30	25	23	-1	1	-23	23
30–40	35 → A	22	0	0	0	0
40–50	45	25	1	1	25	25
50–60	55	10	2	4	20	40
60–70	65	5	3	9	15	45
70–80	75	10	4	16	40	160
		$\Sigma f=125$			$\Sigma fd=2$	$\Sigma fd^2=488$

We know that

$$\sigma=\sqrt{\frac{\Sigma fd^2}{\Sigma f}-\left(\frac{\Sigma fd}{\Sigma f}\right)^2}\times h$$

$$\sigma=\sqrt{\frac{488}{125}-\left(\frac{2}{125}\right)^2}\times 10$$

$$\sigma=\sqrt{3.904-0.0003}\times 10$$

$$\sigma=1.976\times 10=19.76$$

$$\boxed{\sigma=19.76}$$

$$\bar{x}=A+\left(\frac{\Sigma fd}{\Sigma f}\right)\times h=35+\frac{2}{125}\times 10$$

$$=35+0.16=35.16$$

$$\boxed{\bar{x}=35.16}$$

$\therefore$ Coefficient of S.D. $=\frac{\sigma}{\bar{x}}=\frac{19.76}{35.16}$

$$\boxed{\text{Coefficient of S.D.}=0.5620}$$

Variance $(\sigma^2) = (19.76)^2 = 390.15$

Coefficient of variation (CV) $= \frac{\sigma}{\bar{x}} \times 100$

$= 0.5620 \times 100$

Coefficient of variation (CV) $= 56.20\%$ Ans

Imp.

Comparison related question-:

Question-: The following data was recieved by testing 60 bulbs of two different Companies:

Length of life (in hours)	Company A sample	Company B sample
700—900	18	6
900—1100	16	42
1100—1300	26	12
Total	60	60

Calculate S.D. and Coefficient of variation and also state which Company's bulb are more uniform ?.

Solution-: Coefficient of variation → for Comparison

CV. ↑ ⇒ series ↓ (less stable)

less stable / Consistant / uniform.

CV ↓ ⇒ series ↑ (more stable)

more consistant / uniform.

Company (A) -:

Class interval	x (mid)	f	$d=\frac{x-1000}{200}$	d^2	fd	fd^2
700-900	800	18	-1	1	-18	18
900-1100	1000 A	16	0	0	0	0
1100-1300	1200	26	1	1	26	26
		$\Sigma f=60$			$\Sigma fd=8$	$\Sigma fd^2=44$

$$\text{Coefficient of variation} = \frac{\sigma}{\bar{x}} \times 100$$

$$\bar{x} = A + \left(\frac{\Sigma fd}{\Sigma f}\right) \times h$$

$$d = \frac{x-A}{h}$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2} \times h$$

$$h = 200$$

$$\sigma = \sqrt{\frac{44}{60} - \left(\frac{8}{60}\right)^2} \times 200$$

$$\sigma = \sqrt{0.73 - 0.02} \times 200$$

$$\sigma = \sqrt{0.71} \times 200$$

$$\sigma = 0.8426 \times 200$$

$$\boxed{\sigma = 168.52}$$

$$\bar{x} = 1000 + \frac{8}{60} \times 200$$

$$\boxed{\bar{x} = 1026.67}$$

Now $CV = \frac{\sigma}{\bar{x}} \times 100$

$$= \frac{168.52}{1026.67} \times 100$$

$$= 16.41\ \%$$

$\boxed{\text{Coefficient of variation (CV)} = 16.41\%}$ (More ⇓ less uniform.)

Company (B)

Class interval	x(mid)	f	$d = \frac{x-1000}{200}$	d^2	fd	fd^2
700–900	800	6	−1	1	−6	6
900–1100	1000 A	42	0	0	0	0
1100–1300	1200	12	1	1	12	12
		$\Sigma f = 60$			$\Sigma fd = 6$	$\Sigma fd^2 = 18$

We know that

$$\sigma = \sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2} \times h \qquad d = \frac{x-A}{h}$$

$$\bar{x} = A + \left(\frac{\Sigma fd}{\Sigma f}\right) \times h$$

$$\text{Coefficient of variation (CV)} = \frac{\sigma}{\bar{x}} \times 100$$

$$\sigma = \sqrt{\frac{18}{60} - \left(\frac{6}{60}\right)^2} \times 200$$

$$\boxed{\sigma = 107.7}$$

$$\bar{x} = 1000 + \left(\frac{6}{60}\right) \times 200$$

$$\boxed{\bar{x} = 1020}$$

Now Coefficient of variation (CV) $= \frac{6}{\bar{x}} \times 100$

$$= \frac{107.7}{1020} \times 100$$

$$CV = 10.56\% \text{ (less)}$$

$\Downarrow$

More uniform.

$\Rightarrow$ Company B bulb is more uniform.