4.1 Measures of Dispersion and Variation

It is help to interpret the variability of data. To know how much homogenous or nonhomogeneous the data

There are two main types of measures of dispersion in statistics which are :

a) Absolut measures of dispersion and variation

b) Relative measures of dispersion and variation

a) Absolut measures of dispersion and variation

1. Range

The range is the difference between the highest and lowest value in a set of data.

- ✓ **It is the simplest measure of dispersion**
- ✓ **It is very sensitive to the smallest and largest data values**
- ✓ **For unclassified data**

R = $\boldsymbol{max_v - min_v}$

Example 4.1 **Find the range of the values 10, 13, 15, 28, 36, 57, 80.**

Solution

R = $\boldsymbol{max_v - min_v}$

= 80-10

= 70

Example 4.2 **Find the range for the following data: 16, 7, 21, 12, 19, 20, and 11.**

Solution

R = $\boldsymbol{max_v - min_v}$

=21-7

= 14

- ✓ **For classified data**

R = L*M

Example 4.3 **Calculate the range for the following data showing weights in kilograms of 100 students:**

Wight (Kg)	Number of students
60-62	5
63-65	40
66-68	28
69-71	15
72-75	12

Solution:

R = L*M

= 3*5

= 15

2. Variance: Is a statistical measurement that is used to determine the spread of the numbers in a data set with respect to the average value or the mean.

- There are two types of variance in statistics, namely, sample variance and population variance.
- The variance is always calculated with respect to the sample mean.
- **The symbol of sample variance is S^2, and for population variance is σ^2.**

✓ **For unclassified data**

A. **sample variance (S^2)**

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

B. population variance is σ^2

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

Example 4. 4 suppose the data set as following (3, 5, 8, 1), find the population variance.

Solution

$$\sigma^2 = \frac{\sum(x_i - \overline{x})^2}{n}$$

- ✓ **Find the mean**
- ✓ $\overline{X} = \frac{\sum x}{n} = \frac{3+5+8+1}{4} = 4.25$

$$\sigma^2 = \frac{(3-4.25)^2+(5-4.25)^2+(8-4.25)^2+(1-4.25)^2}{4}$$

= 6.68

Example 4. 5 Calculate the sample variance for data representing 8 students' grades:

7, 6, 10, 9, 8, 5, 5, 6.

Solution

$$S^2 = \frac{\sum(x_i - \overline{x})^2}{n-1}$$

- ✓ $\overline{X} = \frac{\sum x_i}{n} = \frac{7+6+10+8+9+5+6+5}{8} = 7$

$$S^2 = \frac{(7-7)^2+(6-7)^2+(10-7)^2+(9-7)^2+(8-7)^2+(5-7)^2+(5-7)^2+(6-7)^2}{8-1}$$

= 3.43

- ✓ **For classified data:**
- ✓ **sample variance (S^2)**

$$S^2 = \frac{\sum f_i(y_i - \overline{x})^2}{\sum f_i - 1}$$

Example 4.6 **Calculate the sample variance for the following data**

Classes	Frequency
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
	$\sum f_i = 100$

Solution

Classes	Frequency	y_i	$f_i * y_i$	$y_i - \bar{x}$	$(y_i - \bar{x})^2$	$f_i(y_i - \bar{x})^2$
60-62	5	61	305	-6.45	41.6	208
63-65	18	64	1152	-3.45	11.9	214.2
66-68	42	67	2814	-0.45	0.203	8.53
69-71	27	70	1890	2.55	6.5	175.5
72-74	8	73	584	5.55	30.8	246.42
	$\sum f_i = 100$		$\sum f_i * y_i = 6745$			$\sum f_i(y_i - \bar{x})^2 =$ 852.65

$$\bar{X} = \frac{\sum f_i * y_i}{\sum f_i}$$

$$\bar{X} = \frac{5*61+18*64+42*67+27*70+8*73}{100}$$

$\bar{X} = 67.45$

$$S^2 = \frac{\sum f_i (y_i - \bar{x})^2}{\sum f_i - 1}$$

$$= \frac{5*(61-67.45)^2+18*(64-67.45)^2+42*(67-67.45)^2+27*(70-67.45)^2+8*(73-67.45)^2}{99}$$

= 8.6

3. Standard Deviation

Is the positive square root of the variance.

- ✓ **For unclassified data**

A. **sample Standard Deviation (S)**

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

B. population Standard Deviation (σ)

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

Example 4.7 **Calculate the population standard deviation of the data mentioned in an example 4.**

Solution

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

$= \sqrt{6.68} = 2.58$

- ✓ **For classified data:**
- ✓ **sample Standard Deviation (S^2)**

$$S = \sqrt{S^2} = \sqrt{\frac{\sum f_i (y_i - \overline{x})^2}{\sum f_i - 1}}$$

Example ٤.8 **Calculate the sample standard deviation of the data mentioned in an example 7.**

Solution

$S = \sqrt{S^2}$

$= \sqrt{8.6} = 2.93$

H.W 4.1 For the following data: 128, 219, 316, 189, 512, 155, 98, 110, 468, 177, 203, 73. Find: the range, the Population variance and standard deviation.

b) Relative measures of dispersion and variation

- ✓ Relative measures of dispersion are important when comparing the dispersion of two groups (or materials), such as comparing the dispersion of lengths to the dispersion of weights.
- ✓ These measures are unitless.
- ✓ One of these measures is the Coefficient of Variation (C.V).

❖ **Coefficient of Variation (C.V)**

- ✓ Is the percentage of the standard deviation (S or σ) and the arithmetic mean ($\bar{X}$).
- ✓ **For the sample**

C.V % = $\frac{S}{\bar{X}}$ * 100 (for unclassified, and classified data).

- ✓ **For the population**

C.V % = $\frac{\sigma}{\bar{X}}$ * 100 (for unclassified, and classified data).

Example 4.9 **Compare the dispersion of the students' lengths and their weights, using the following data, which represent the lengths and weights of 100 students.**

Lengths (cm)	Number of students
110 – 119	12
120 – 129	15
130 – 139	25
140 – 149	30
150 – 159	10
160 - 170	8
	$\sum f_i$= 100

Weights (Kg)	Number of students
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
	$\sum f_i$= 100

Solution

A. **Dispersion of weight**

Weights (Kg)	Number of students	y_i
60-62	5	61
63-65	18	64
66-68	42	67
69-71	27	70
72-74	8	73
	$\sum f_i$= 100	

$$\bar{X} = \frac{5*61+18*64+42*67+27*70+8*73}{100}$$

$\bar{X}$ = 67.45

$$S^2 = \frac{\sum f_i(y_i-\bar{x})^2}{\sum f_i - 1}$$

$$= \frac{5*(61-67.45)^2+18*(64-67.45)^2+42*(67-67.45)^2+27*(70-67.45)^2+8*(73-67.45)^2}{99}$$

= 8.6

$S = \sqrt{S^2} \rightarrow \sqrt{8.6}$ = 2.93

$$C.V\ \% = \frac{S}{\bar{X}} * 100$$

$$= \frac{2.93}{67.45} * 100 \rightarrow 4.34\%$$

B. Dispersion of length

Lengths (cm)	Number of students	y_i
110 – 119	12	114.5
120 – 129	15	124.5
130 – 139	25	134.5
140 – 149	30	144.5
150 – 159	10	154.5
160 - 170	8	165
	$\sum f_i$= 100	

$$\bar{X} = \frac{12*114.5+15*124.5+25*134.5+30*144.5+10*154.5+8*165}{100}$$

$\bar{X}$ = 138.04

$$S^2 = \frac{\sum f_i(y_i-\bar{x})^2}{\sum f_i - 1}$$

$$\frac{12*(114.5-138.04)^2+15*(124.5-138.04)^2+25*(134.5-138.04)^2+30*(144.5-138.04)^2+10(154.5-138.04)^2+8(165-138.04)^2}{99}$$

= 196.86

$S = \sqrt{S^2} \rightarrow \sqrt{196.86}$ = 14.03

$$C.V\ \% = \frac{S}{\bar{X}} * 100$$

$$= \frac{14.03}{138.04} * 100 \rightarrow 10.19\ \%$$

Dispersion of length is more than Dispersion of weight.