Practical.4

Measures of dispersion – variance, standard deviation and coefficient of variation for raw data

Variance

The square of the standard deviation is called variance

(i.e) variance = $\sqrt{SD}$

Standard Deviation

It is defined as the positive square-root of the arithmetic mean of the Square of the deviations of the given observation from their arithmetic mean.

The standard deviation is denoted by S in case of sample and Greek letter s (sigma) in case of population.

The formula for calculating standard deviation is as follows

$$S = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}}$$ for raw data

And for grouped data the formulas are

$$S = \sqrt{\frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N}\right)^2}$$ for discrete data\

$$S = C \times \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$$ for continuous data

Where d = $\frac{x - A}{C}$

C = class interval

Example 1

Raw Data

The weights of 5 ear-heads of sorghum are 100, 102,118,124,126 gms. Find the standard deviation.

Solution

	x	**x^2**
	100	10000
	102	10404
	118	13924
	124	15376
	126	15876
Total	570	65580

Standard deviation $S = \sqrt{\dfrac{\sum x^2 - \dfrac{(\sum x)^2}{n}}{n-1}}$

$$= \sqrt{\frac{65580 - \dfrac{(570)^2}{5}}{5-1}} = \sqrt{150} = 12.25 \; gms$$

Variance = $\sqrt{12.25}$ =3.5

Coefficient of Variation

The Standard deviation is an absolute measure of dispersion. It is expressed in terms of units in which the original figures are collected and stated. The standard deviation of heights of students cannot be compared with the standard deviation of weights of students, as both are expressed in different units, i.e heights in centimeter and weights in kilograms. Therefore the standard deviation must be converted into a relative measure of dispersion for the purpose of comparison. The relative measure is known as

the coefficient of variation. The coefficient of variation is obtained by dividing the standard deviation by the mean and expressed in percentage. Symbolically, Coefficient of variation (C.V) = $\frac{SD}{mean} \times 100$

If we want to compare the variability of two or more series, we can use C.V. The series or groups of data for which the C.V. is greater indicate that the group is more variable, less stable, less uniform, less consistent or less homogeneous. If the C.V. is less, it indicates that the group is less variable or more stable or more uniform or more consistent or more homogeneous.

Example 2

Consider the measurement on yield and plant height of a paddy variety. The mean and standard deviation for yield are 50 kg and 10 kg respectively. The mean and standard deviation for plant height are 55 am and 5 cm respectively.

Here the measurements for yield and plant height are in different units. Hence the variabilities can be compared only by using coefficient of variation.

For yield, CV= $\frac{10}{50} \times 100 = 20\%$

For plant height, CV= $\frac{5}{55} \times 100 = 9.1\%$

The yield is subject to more variation than the plant height.

Learning Exercise

1.	The weights of 8 earheads of sorghum are 14, 29, 9, 15, 20, 17, 12, and 11. Find Standard Deviation and Variance and coefficient of variation.
2.	Find out which of the following batsmen is more consistent in scoring. Batsman A 5 7 16 27 39 53 56 61 80 101 105 Batsman B 0 4 16 21 41 43 57 78 83 93 95