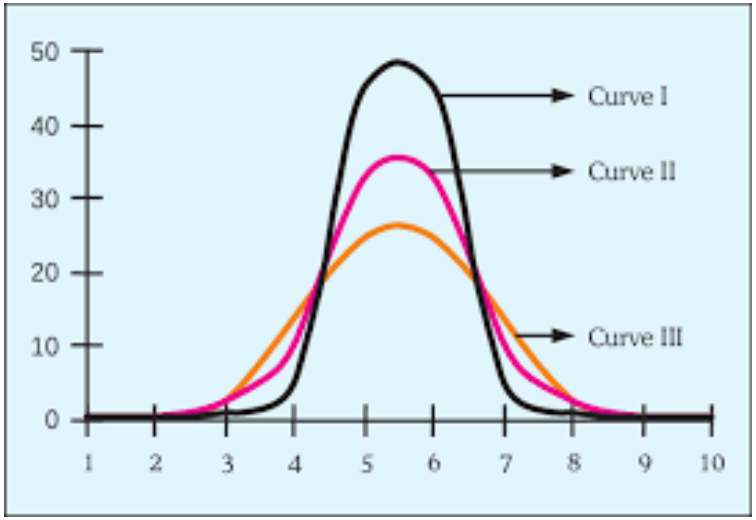
Measures of Dispersion

Ansar Shahzadi
School of Electrical Engineering & Computer Science
National University of Science and Technology(NUST)

Why study Dispersion?

- An average just describes the center of the data.
- Only knowing averages, we are therefore not able to tell how the data dispersed about the average?



What is Dispersion

Dispersion(also known scatter, spread, or Variation) measures the items vary from some central value. Its measures the degree of variation.

- The more similar the scores are to each other, the lower the measure of dispersion will be
- The less similar the scores are to each other the higher the measure of dispersion will be

Significance of Measuring Dispersion

- To describe about the spread of data.
- Its uses the basis of comparisons.
- To maintain standard.
- To facilitate the use of other statistical measures.

Measures of Dispersion

- Range
- Quartile Deviation
- Standard Deviation
- Variance
- Coefficient of Variation

Range

Difference between the largest and smallest value of the data

$$Range = x_{max} - x_{min}$$

Quartile Deviation

- **Interquartile Range…** Difference between the upper and lower quartile of the data

- **Quartile Deviation…** half of the semi-interquartile range

$$Q.D = \frac{Q_3 - Q_1}{2}$$

Question 1

The frequency distribution of 160 determinations of the daily emission (in tons) of sulfur oxides from an industrial plant is given below. Find the quartile deviation.

Class Limits	Frequency
5-8.9	8
9-12.9	20
13-16.9	29
17-20.9	45
21-24.9	32
25-28.9	19
29-32.9	7
Total	160

Standard Deviation

Standard Deviation is defined as the positive squared root of the average of the square deviations.

For ungrouped data

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

For grouped data

$$s=\sqrt{\frac{\sum f_i(x_i-\bar{x})^2}{\sum f_i}}$$

Variance

Variance is defined as the average of the square deviations.

For ungrouped data

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

For grouped data

$$s^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$$

Question 2

The frequency distribution of 160 determinations of the daily emission (in tons) of sulfur oxides from an industrial plant is given below. Find the Standard deviation and Variance.

Class Limits	Frequency
5-8.9	8
9-12.9	20
13-16.9	29
17-20.9	45
21-24.9	32
25-28.9	19
29-32.9	7
Total	160

Properties of Standard Deviation

- S.D(a)= o
- S.D(X+a)=S.D(X)
- S.D(ax)= $|a|$S.D(x)
- If x and Y are two independent variables, then S.D(X±Y)=$\sqrt{\text{Var(X)+Var(Y)}}$
- $s_c = \sqrt{\frac{1}{\sum n_i}\sum n_i \left(s_i^2 + (\bar{x}_i - \bar{x}_c)^2\right)}$

Question 3

- For the following data

weights	74.5	94.5	114.5	134.5	154.5	174.5	194.5
frequency	9	10	17	17	5	4	5

Calculate the Mean and standard deviation using

- X
- D=(X-134.5)
- U=(X-74.5)/20

And discuss the result.

Coefficient of variation

The coefficient of variation is obtained by dividing the standard deviation by the mean and expressed in percentage. Symbolically,

$$c.v = \frac{standard\ deviation}{mean} \times 100$$

If we want to compare the variability of two or more series, we can use C.V.

- If the C.V. is greater indicates 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.