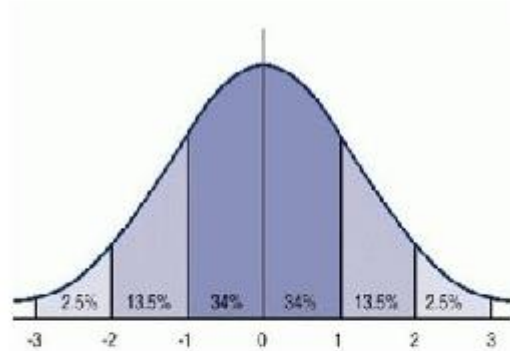# STATISTICAL FUNDAMENTALS

Course : Total Quality Management

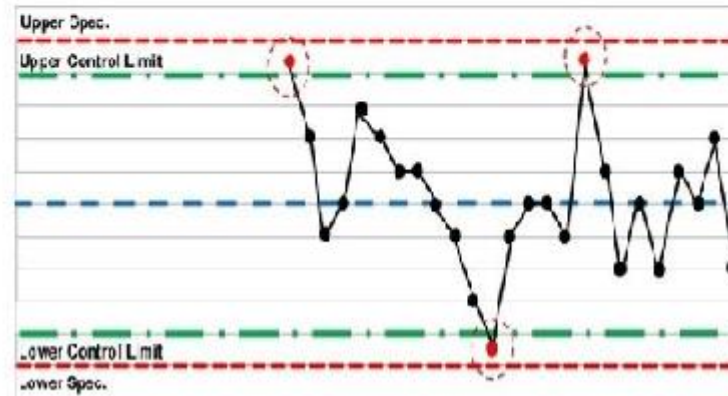**19ME514**
Unit -3 Statistical Process Control
III Year /V Semester
Mechanical Engineering

TOPIC OF THE DAY

STATISTICAL FUNDAMENTALS

Statistics is defined as the science that deals with the collection, tabulation, analysis, interpretation and presentation of quantitative data.

Data collected for quality control purposes are obtained by direct observation and are classified as

1. Variables (Measurable quality characteristics like length measured in metres)

2. Attributes (Quality characteristic which are classified as either conforming (or) non-conforming to specifications, such as "go & no-go" gauge.

MEASURES OF CENTRAL TENDENCY AND DISPERSION

There are two important analytical methods of describing a collection of data as

1. Measures of central tendency.

2. Measures of dispersion.

A measure of central tendency of a distribution is a numerical value that describes

how the data tend to build up in the centre. There are three measures in quality as

1. Average

2. Median

3. Mode

AVERAGE

It is the sum of observations divided by the number of observations.

$$\text{Average} = \overline{X} = \frac{\displaystyle\sum_{i=1}^{i=n} X_i}{n}$$

where, n = number of observations

X_i = observed value

MEDIAN AND MODE

Median is the value which divides a series of ordered observations so that the number of items above it is equal to the number of items below it.

Mode is the value which occurs with the greatest frequency in a set of numbers.

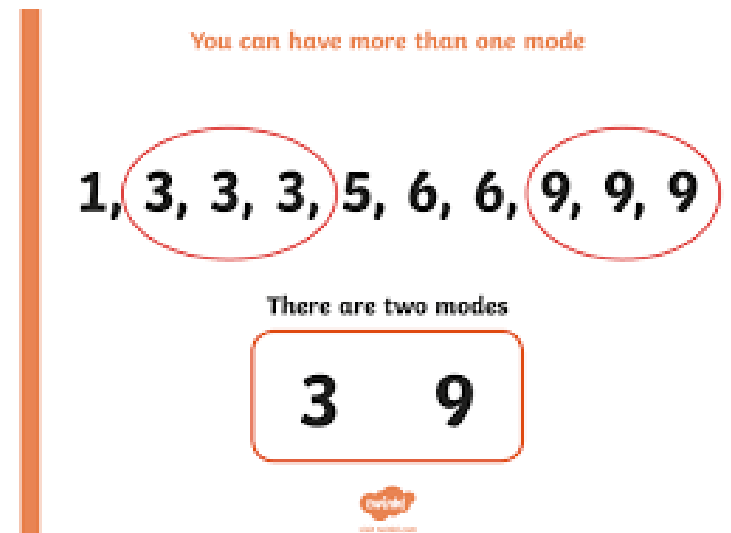1, 3, 3, **6**, 7, 8, 9

Median = 6

1, 2, 3, **4, 5**, 6, 8, 9

Median = (4 + 5) ÷ 2

= 4.5

You can have more than one mode

1, 3, 3, 3, 5, 6, 6, 9, 9, 9

There are two modes

3   9

MEASURE OF DISPERSION

Measure of dispersion describes how the data are spread out on each side of the central value.

The two measures of dispersion are

1. Range

2. Standard Deviation

Range is the difference between the largest and smallest values of observations in a series of numbers.

Range = R = $X_h - X_l$

Where, R = Range

X_h = highest observation in a series

X_l = lowest observation in a series

STANDARD DEVIATION

Standard Deviation measures the spreading tendency of the data. Larger the standard deviation, greater the variability of data.

$$S = \sqrt{\frac{\sum_{i=1}^{i=n} (X_i - \overline{X})^2}{n-1}}$$

where

- S = sample standard deviation
- X_i = observed value
- n = number of observations

POPULATION AND SAMPLE

In order to construct a frequency distribution of the outer diameter of shafts, a small portion (or) sample is selected to represent all the shafts. The population is the whole collection of shafts.

The population may be an hour's production, a week's production, 10000 pieces and so on. It is not possible to measure all of the population. Hence, we go for sampling.

Sampling becomes necessary

1. When it is impossible to measure the entire population.

2. When it is more expensive to observe all the data.

3. When the required inspection destroys the product.

4. When a test of the entire population may be too dangerous as in the case of new medical drug.
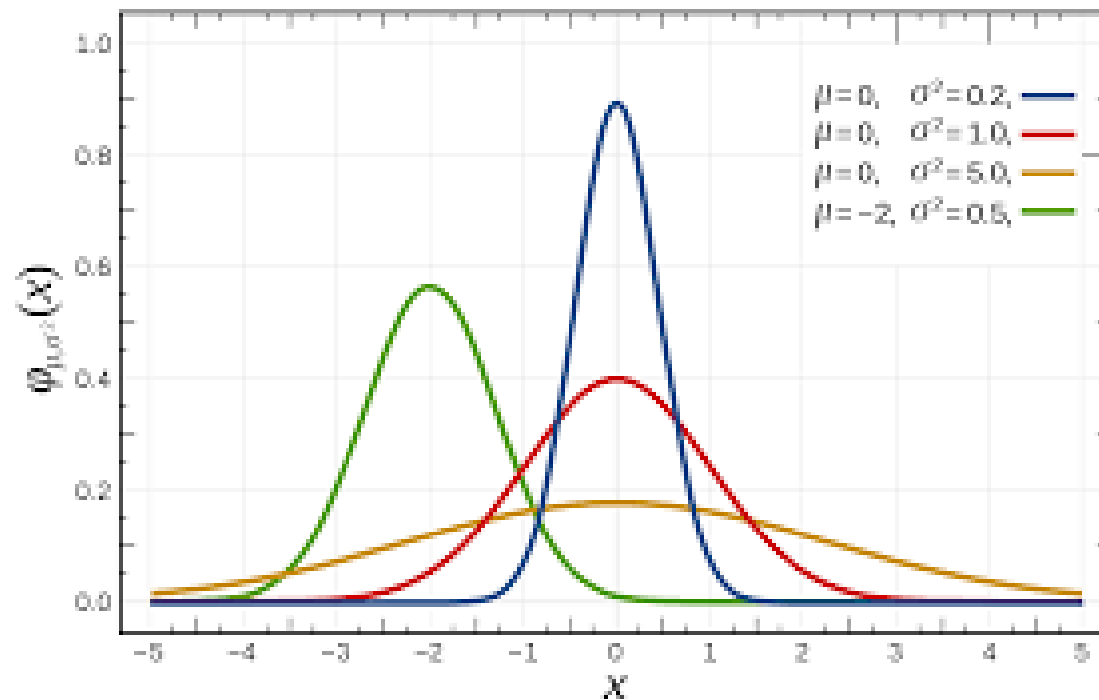
X is for sample average or sample mean.

μ is for population mean.

S is for sample standard deviation.

σ is for population standard deviation.

NORMAL CURVE

Normal curve is common type of population. The normal curve is symmetrical, unimodal, bell – shaped distribution with the mean, median and mode all having the same value.

REFERENCES

1. Dale H. Besterfiled, "Total Quality Management", Third Edition, Pearson Education Asia, Indian Reprint, 2016.

2. Dr S. Kumar, "Total Quality Management", Laxmi Publications Ltd., New Delhi 2006.