

Chemical Process Instrumentation
Prof. Debasis Sarkar
Department of Chemical Engineering
Indian Institute of Technology, Kharagpur

Lecture – 12
Performance Characteristics of instruments and Data Analysis-II (Contd.)

Welcome to lecture 12. So, we are now talking about performance characteristics of instruments and data analysis part 2. So, in this lecture, we will talk about types of different errors that are associated with your measurements. So, today's topic types of error in measurement and also we will introduce some basics statistical concepts.

(Refer Slide Time: 00:31)

Performance Characteristics of Instruments and Data Analysis - 2

Today's Topic:

- ✓ Types of error in measurement
- ✓ Some basic statistical concepts

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

(Refer Slide Time: 00:39)

Types of Error In Measurement

Types of error in measurement:

1. **Gross error/Human error**
2. **Systematic/Determinate Error**
3. **Random Error**

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

The slide features a yellow background with a red title. A list of three error types is presented in a numbered format. The first and third items are in black, while the second is in red. A small video inset of the presenter is visible in the bottom right corner.

Types of errors in measurements can be broadly categorized in 3 groups gross error or human error systematic error or determinate error and random error.

(Refer Slide Time: 00:58).

Types of Error In Measurement

Gross error/Human error

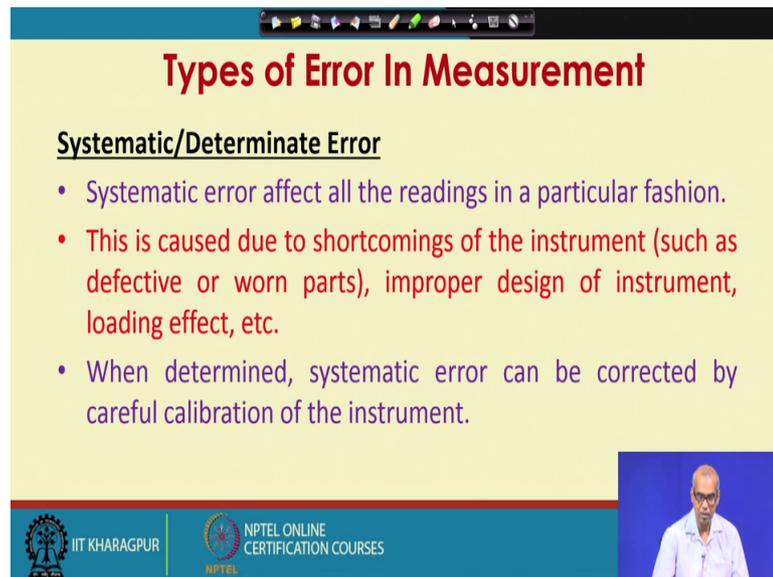
- This is caused by human mistakes: Wrong reading, bad using instruments.
- This cannot be eliminated completely, but can be minimized by taking care and proper usage of instruments.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

The slide features a yellow background with a red title. The section title 'Gross error/Human error' is underlined. Two bullet points describe the error type. The first is in black, and the second is in red. A small video inset of the presenter is visible in the bottom right corner.

Gross error or human error is caused by human mistakes this may happened due to wrong reading bad or wrong using of instruments.

(Refer Slide Time: 01:19)



The slide is titled "Types of Error In Measurement" in red text. Below the title, the section "Systematic/Determinate Error" is underlined. It contains three bullet points: 1. Systematic error affect all the readings in a particular fashion. 2. This is caused due to shortcomings of the instrument (such as defective or worn parts), improper design of instrument, loading effect, etc. 3. When determined, systematic error can be corrected by careful calibration of the instrument. The slide footer includes the IIT KHARAGPUR logo, the NPTEL ONLINE CERTIFICATION COURSES logo, and a small video inset of a speaker.

Types of Error In Measurement

Systematic/Determinate Error

- Systematic error affect all the readings in a particular fashion.
- This is caused due to shortcomings of the instrument (such as defective or worn parts), improper design of instrument, loading effect, etc.
- When determined, systematic error can be corrected by careful calibration of the instrument.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

This cannot be eliminated completely, but can be minimized by taking care and proper usage of instruments systematic or determinate error; systematic error effect all the readings in a particular fashion.

So, they are systematic in nature they are also deterministic that is why determinate this is caused due to shortcomings of the instruments such as defective or worn parts improper design of instrument loading effect etcetera when determined systematic error can be corrected by careful calibration of the instrument that is why you call it determinate error.

(Refer Slide Time: 01:59)

Types of Error In Measurement

Systematic/Determinate Error

- Systematic error may be of a constant or proportional nature

• **Proportional error** influences the slope

• **Constant error** influences the intercept

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Systematic or determinate error may be of a constant or proportional nature when it is proportional in nature it influences the slope constant error influences the intercept.

(Refer Slide Time: 02:34)

Types of Error In Measurement

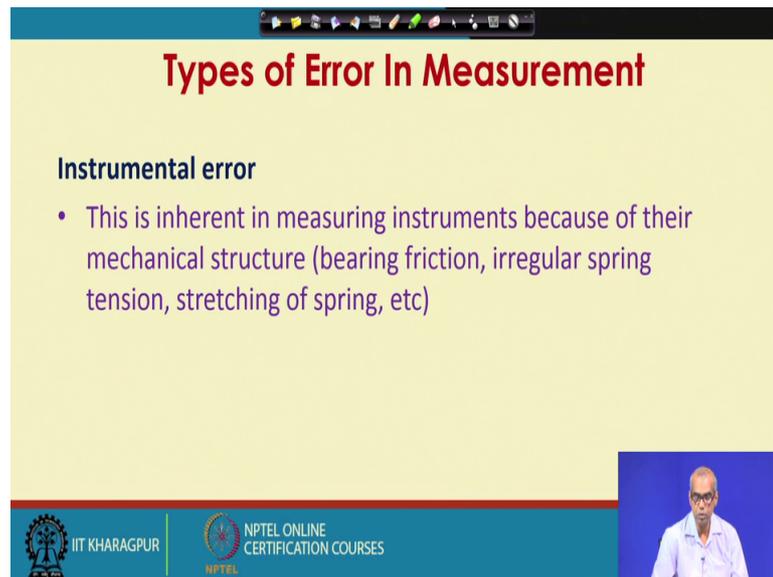
Systematic/Determinate Error

- Mainly THREE types of systematic error :
 - I. *Instrumental error*
 - II. *Environmental error*
 - III. *Observational error*

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Systematic or determinate error are mainly 3 types; instrumental error environmental error and observational error.

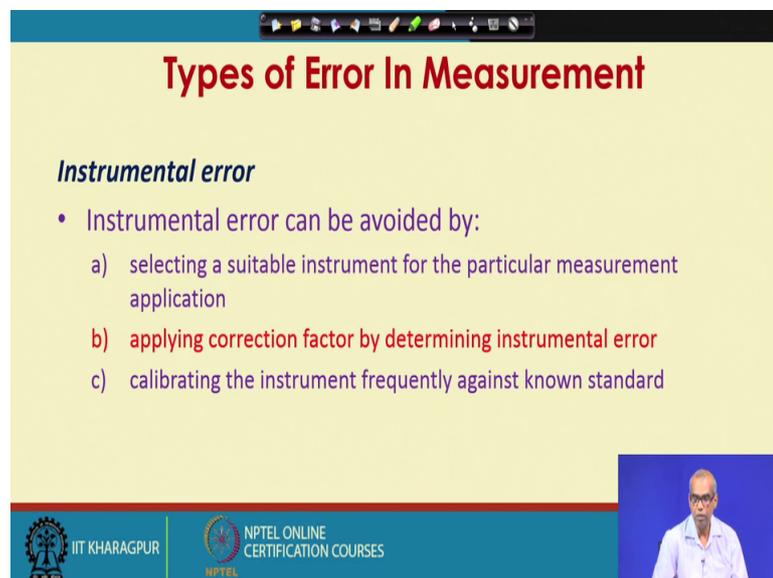
(Refer Slide Time: 02:48)



The slide features a yellow background with a red title 'Types of Error In Measurement'. Below the title, the text 'Instrumental error' is followed by a bullet point: 'This is inherent in measuring instruments because of their mechanical structure (bearing friction, irregular spring tension, stretching of spring, etc)'. The slide includes logos for IIT KHARAGPUR and NPTEL ONLINE CERTIFICATION COURSES, and a small video inset of a speaker in the bottom right corner.

Instrumental error is inherent in measuring instruments because of their mechanical structure such as bearing friction irregular spring tension stretching of spring etcetera.

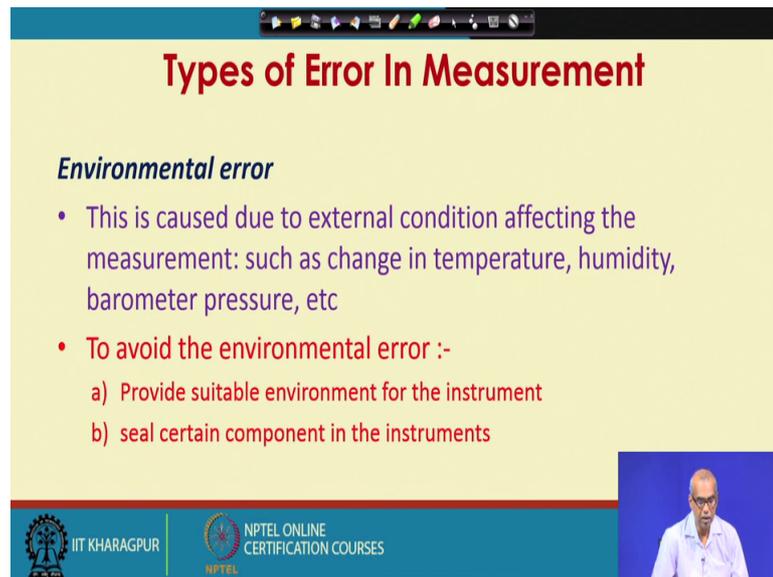
(Refer Slide Time: 03:01)



The slide features a yellow background with a red title 'Types of Error In Measurement'. Below the title, the text 'Instrumental error' is followed by a bullet point: 'Instrumental error can be avoided by:'. This is followed by three sub-points: 'a) selecting a suitable instrument for the particular measurement application', 'b) applying correction factor by determining instrumental error', and 'c) calibrating the instrument frequently against known standard'. The slide includes logos for IIT KHARAGPUR and NPTEL ONLINE CERTIFICATION COURSES, and a small video inset of a speaker in the bottom right corner.

Instrumental error can be avoided by selecting a suitable instrument for the particular measurement application applying a suitable correction factor by determining instrumental error and calibrating the instrument frequently against known standard.

(Refer Slide Time: 03:24)



Types of Error In Measurement

Environmental error

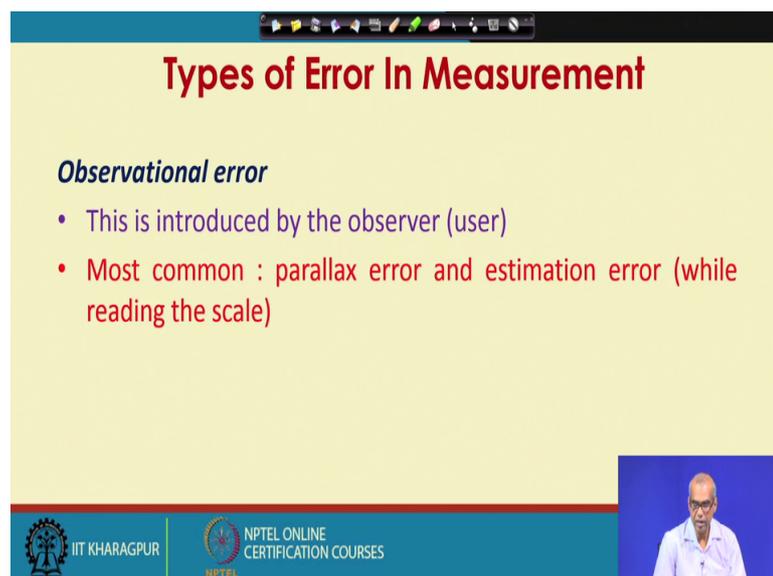
- This is caused due to external condition affecting the measurement: such as change in temperature, humidity, barometer pressure, etc
- To avoid the environmental error :-
 - a) Provide suitable environment for the instrument
 - b) seal certain component in the instruments

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES



Environmental error is caused due to external condition effecting the measurement such as change in temperature humidity barometric pressure, etcetera to avoid the environmental error provide suitable environment for the instrument seal certain component in the instruments; that means, if some component of the instruments are sensitive source temperature we put that component in a sealed environment to isolate it from the effect of environment.

(Refer Slide Time: 04:05)



Types of Error In Measurement

Observational error

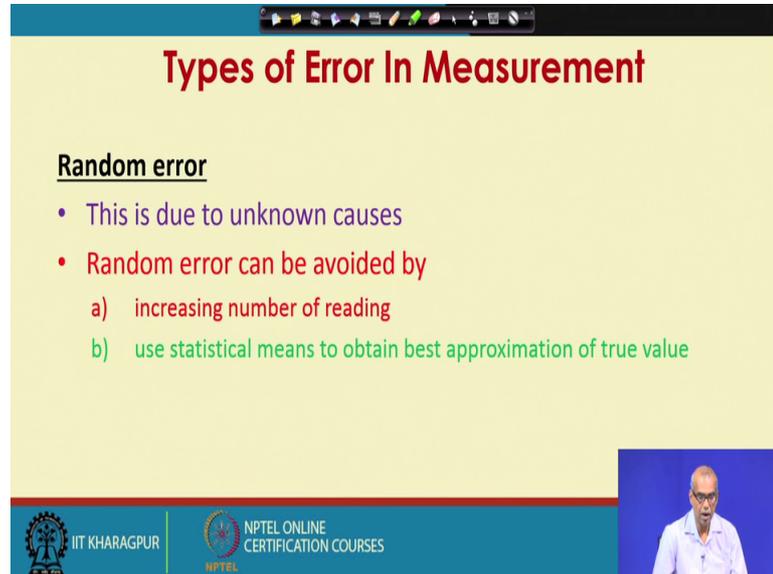
- This is introduced by the observer (user)
- Most common : parallax error and estimation error (while reading the scale)

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES



Observational error; this is introduced by the observer or the user or the operator of the instrument most common is parallax error and estimation error while reading the scale.

(Refer Slide Time: 04:21)



Types of Error In Measurement

Random error

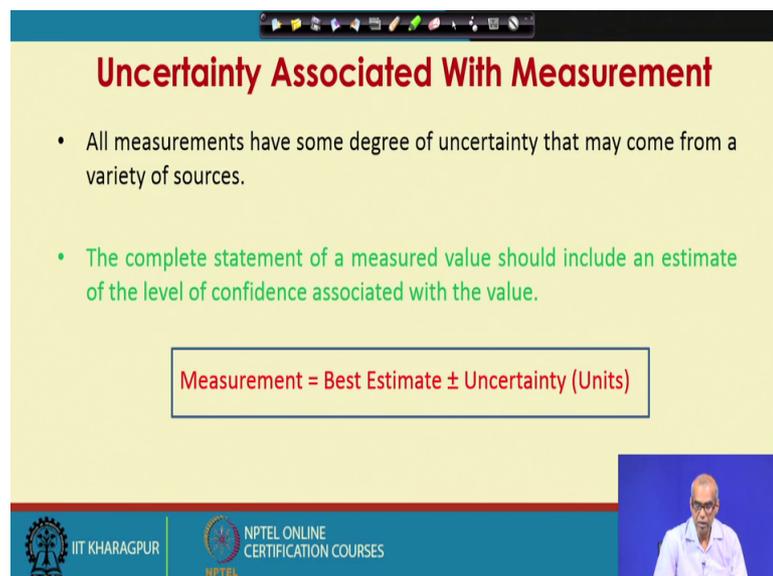
- This is due to unknown causes
- Random error can be avoided by
 - a) increasing number of reading
 - b) use statistical means to obtain best approximation of true value

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Random error is due to unknown causes random error can be avoided by increasing number of reading.

So, you take more number of readings and use statistical means to obtain best approximation of true value.

(Refer Slide Time: 04:42)



Uncertainty Associated With Measurement

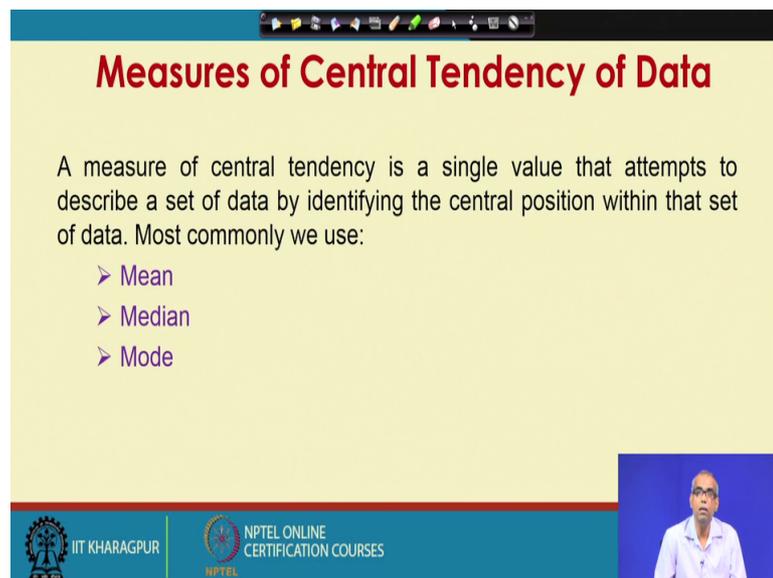
- All measurements have some degree of uncertainty that may come from a variety of sources.
- The complete statement of a measured value should include an estimate of the level of confidence associated with the value.

Measurement = Best Estimate \pm Uncertainty (Units)

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Now, let us talk about uncertainty associated with measurement all measurements have some degree of uncertainty that may have come from variety of sources the complete statement of a measured value should include an estimate of the level of confidence associated with the value.

(Refer Slide Time: 05:28)



The slide features a yellow background with a red title 'Measures of Central Tendency of Data'. Below the title, a paragraph defines a measure of central tendency as a single value representing the central position of a data set. A bulleted list follows, listing 'Mean', 'Median', and 'Mode'. The slide is part of an NPTEL presentation, as indicated by the logos for IIT Kharagpur and NPTEL Online Certification Courses at the bottom. A small video inset in the bottom right corner shows a male presenter.

So, measurement is best estimate plus minus some uncertainty measures of central tendency of data a measure of central tendency is a single value that attempts to describe a set of data by identifying the central position within that set of data imagine; you have taken an instrument and you are measuring a quantity a several times.

So, you have a set of data a measure of central tendency is a single value that attempts to describe a set of data by identifying the central position within that set of data most commonly we use mean may be arithmetic mean median and mode. So, we will now define what do you mean by mean median and mode.

(Refer Slide Time: 06:42)

Measures of Central Tendency of Data

Arithmetic Mean: The mean is equal to the sum of all the values in the data set divided by the number of values in the data set.

$$\bar{X} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{\sum_{i=1}^N x_i}{N}$$

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Arithmetic mean is equal to the sum of all the values in the data set divided by the number of values in the data set. So, I have taken n observations or n readings those values are x_1, x_2, \dots, x_n . So, the arithmetic mean is simply sum of all these values sum of n readings divided by n which represent the number of readings taken. So, arithmetic mean is can be written as x_i i equal to 1 to n here and n.

(Refer Slide Time: 07:55)

Measures of Central Tendency of Data

Arithmetic Mean: The mean is equal to the sum of all the values in the data set divided by the number of values in the data set.

$$\bar{X} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N}$$

Weighted Mean:
$$\bar{X} = \frac{w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_Nx_N}{w_1 + w_2 + w_3 + \dots + w_N} = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i}$$

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Now, we can also have weighted mean weighted mean is defined as this.

So, you give certain weightage to each observation and let us consider these weights are w_1, w_2, w_3, w_n , etcetera. So, w_1 is associated with the observation x_1 , w_2 is the weight associated with the observation x_2 and so on and so forth. So, weighted mean is defined as $w_1 x_1 + w_2 x_2 + \dots + w_n x_n$ divided by $w_1 + w_2 + \dots + w_n$. So, you can write as $\sum_{i=1}^n w_i x_i$ divided by $\sum_{i=1}^n w_i$. So, this is weighted mean.

(Refer Slide Time: 09:24)

Measures of Central Tendency of Data

Median: The median is defined as the middle value of a group of data arranged in order of magnitude.

Median = $\{x\}_{(N+1)/2}$ when N is odd ✓ ✓ N = 11 #

Median = $\frac{1}{2}[\{x\}_{N/2} + \{x\}_{(N/2)+1}]$ when N is even ✓ ✓

Handwritten calculations:
 $\frac{11+1}{2} = \frac{12}{2} = 6$
 $\frac{10}{2} = 5$
 $\frac{10+1}{2} = \frac{11}{2} = 5.5$
 $\frac{1}{2}[5+6] = 5.5$

Diagram: A number line with 5 and 6. A vertical line is drawn at 5.5, with a circle around the number 6 below it.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES | 16

Next is median the median is defined as the middle value of a group of data arranged in order of magnitude? So, you have a set of data you have taken let us say you have taken a pressure measuring instrument and you have measured a given pressure several times. So, you have a set of data; now you arrange this data in order; now the median is the middle value of this arranged data.

So, median is defined as the middle value of a group of data arranged in order of magnitude now depending on the number of observation that you have taken you can have an exact midpoint let us say you have taken 11 readings you have taken 11 readings. So, in that case sixth one is the midpoint. So, 5 readings above 6; 5 readings below 6. So, 6 is the midpoint 6 is the median, but let us say, I have taken 10 reading, then I have to take average of 5 and 6 as median. So, when n is odd your median will be given by this and when n is even it will be given by this. So, when my N is 11. So, number of observations is 11 median is $N + 1$ by 2; that means, 11 plus 1 by 2 that is

the 6; data 6 from the top they are all arranged in order of magnitude N is odd. So, you get the sixth one if N is 10, it is even. So, you said that it has to be average of 5 and 6.

So, half N by 2, 10 by 2 plus 10 by 2 plus 1, 10 by 2 plus 1. So, half 5 plus 6. So, it I think there is some problem here. So, it will be N is even. So, 10 it will be four top another four below you have 5 and 6. So, this average of 5 and 6. So, 5 and 6 average of 5 and 6 this fine, no, no problem here this fine average of 5 and 6 average of fifth and sixth value. So, this is correct this is correct we have to calculate median depending on whether the number of reservations is even or odd accordingly mode the mode of a set of data is the value that appears most frequently the mode of a set of data is the value that appears most frequently. Now if I have several observations and then I make a frequency plot like this.

(Refer Slide Time: 13:37)

Measures of Central Tendency of Data

Median: The median is defined as the middle value of a group of data arranged in order of magnitude.

$$\text{Median} = \{x\}_{(N+1)/2} \quad \text{when } N \text{ is odd}$$

$$= \frac{1}{2} [\{x\}_{N/2} + \{x\}_{(N/2)+1}] \quad \text{when } N \text{ is even}$$

Mode: The mode of a set of data is the value that appears most frequently.

(b)

Unimodal

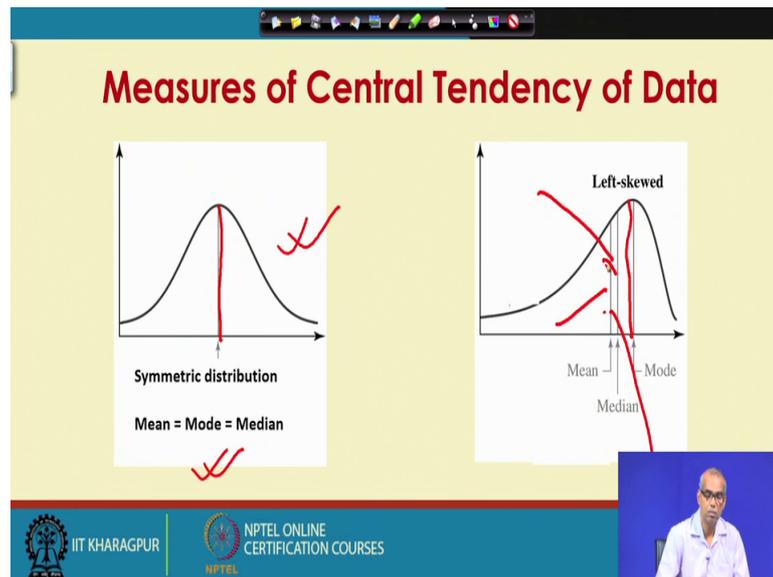
(c)

Bimodal distribution

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES | 16

So, basically you plot observations those values verses frequency and if I get a curve like this; this is the value that happens most frequently. So, this is the mode of this set of data it is not necessary that there will be only one mode if there is only one mode we call it unimodal if the plot is like this frequency verses observation you have one mode here you have another mode here. So, this is known as bimodal distribution.

(Refer Slide Time: 15:16)



If you have a symmetric distribution like this your mean mode and median will all be the same because if you have series in if you have a symmetric distribution this is mean this is mode as well and this is median as well, but if you have instead skewed distribution; that means, the distribution is not symmetric it is skewed it is called left skewed here you can also have right skewed then this is mode.

Because this is the value that you encounter most frequently, but mean may be this and the median this is mean and mode median may be this. So, mean median and mode may all be different mean median and mode are same only for a symmetric distribution, but distributions which are not symmetric the mean median mode may all be different.

(Refer Slide Time: 16:55)

Measures of Dispersion (Spread of Variability)

Range: Range is the difference between the maximum and minimum values of the given data.

- It is based solely on the dispersion of the extreme values.

Average or Mean Deviation: It evaluates the average of the absolute value of the deviation of each data point from the arithmetic mean.

$$MD = \frac{1}{N-1} \sum_{i=1}^N |x_i - \bar{X}|$$

The slide also features logos for IIT KHARAGPUR and NPTEL ONLINE CERTIFICATION COURSES, and a small video inset of a speaker.

Measures of dispersion or spread of variability range is the difference between the maximum and minimum values of the given data we have talked about this one we talked about static characteristics. So, if I take range to discuss the measure of dispersion or the spread of variability that is how the data is distributed or how the data is spread one can use the definition of range which is the difference between the maximum value and minimum value of the given data, but this is based only on the dispersion of the extreme values extreme high and extreme low average or mean deviation.

This is a better estimate of measure of dispersion or spread of variation or spread of variability average or mean deviation evaluates the average of the absolute value of the deviation of each data point from the arithmetic mean. So, average or mean deviation evaluates the average of the absolute value of the deviation of each data point from the arithmetic mean you have N data points. So, you find out the absolute difference from each observation from the mean which is represented as this and sum those differences and then divide by N minus 1 to get average or mean deviation.

(Refer Slide Time: 18:58)

Measures of Dispersion (Spread of Variability)

Variance: It is the mean square deviations of the given data.

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$$

Standard Deviation: It is the square root of the variance.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2}$$

The slide includes logos for IIT KHARAGPUR and NPTEL ONLINE CERTIFICATION COURSES, and a small video inset of a presenter in the bottom right corner.

Variance it is the mean square deviation of the given data look at the expression for variance and at the expression for the average or mean deviation. So, variance is the mean square deviations of the given data.

So, you have find out you have found out the difference of the observation from the mean difference of each observation from mean square them sum them up and divide by N minus 1 where capital N is the number of the observation standard deviation is just the square root of the variance. So, this is the expression for variance square root of this is the expression for standard deviation.

(Refer Slide Time: 20:13)

Measures of Dispersion (Spread of Variability)

Coefficient of Variation: The coefficient of variation represents the degree of scatter of the data with respect to the mean value.

$$C.V = \frac{\sigma}{\bar{X}}$$

It may be noted that the coefficient of variation fails to be useful when mean is close to zero.

CV Low

CV High

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES | 20

Coefficient of variation the coefficient of variation presents the degree of scatter of the data with respect to the mean value. So, the coefficient of variation the coefficient of variation represents the degree of scatter of data with respect to mean value. So, this is ratio of sigma to mean value sigma or standard deviation is given by this. So, you divide sigma by mean value of the data set.

So, you get coefficient of variation please note that the coefficient of variation will not be useful when mean is very close to 0. So, if you have a distribution like this on one hand and if I have a distribution like these on other hand here you see that the spread of the data around me or the scatter of data around mean is low here the spread or scatter of data around mean is high. So, the coefficient of variation is low here the coefficient of variation is high here, here most of the data are centered around mean here most of the data are not centered around mean.

(Refer Slide Time: 22:15)

Estimate of Uncertainty: Standard Error

$X = \bar{X}_n \pm U_n$

$U_n = \pm \frac{s_n}{\sqrt{n}}$

\bar{X}_n is the mean of the sample data
 U_n is the best estimate of internal uncertainty
 s_n is the unbiased estimate of standard deviation

The slide features a yellow background with a blue header and footer. The title is in red. The equations are in black. Red circles and arrows highlight the equations. A small video inset of a speaker is in the bottom right corner.

So, this is how we estimate or represent the uncertainty we use standard error if \bar{X}_n is the mean of the sample data and U_n is the best estimate of internal uncertainty we will represent the measured data as mean plus minus best estimate of uncertainty.

So, X is \bar{X}_n plus minus U_n and an estimate of this uncertainty is standard deviation this is standard deviation divided by square root of n where n is the number of observation. So, an estimate of standard error or uncertainty is standard deviation divided by square root of number of observations and then the measured value will be expressed as mean value plus minus this U_n which is the best estimate of the internal uncertainty can be represented as standard deviation divided by square root of number of observations. So, we will stop our lecture 12 here.