Introduction to Statistics

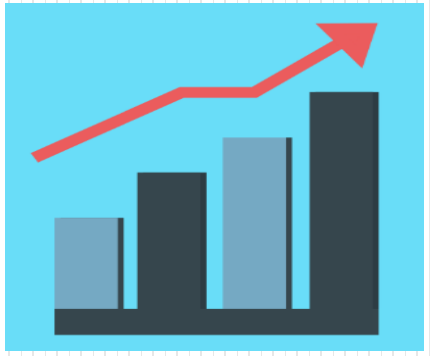
Dr. Farman Ali

Assistant Professor

DEPARTMENT OF SOFTWARE

SEJONG UNIVERSITY

Lecture-3



Course Syllabus

- Introduction to the course
- Sampling and data presentation
- Basic of probability
- Distributions
- Confidence intervals
- Hypothesis testing
- Correlation and simple linear regression
- Multiple regression

Outline

➢ Data Representation

- Classification
- Tabulation
- Graphs

➢ Measures of Central Tendency

- Mean
- Mode
- Median

Organizing Data

- To get an understanding of the data, it is organized and arranged into a meaningful form. This is done by the following methods:

- **Classification**

- **Tabulation** (e.g. simple tables, frequency tables, stem and leaf plots etc.)

- **Graphs** (Bar Graph, Pie chart, Histogram, etc)

Tabulation of Data

- The process of placing classified data into tabular form is known as tabulation.
- To present data in the form of rows and column.

- There are three types of tabulation.

❖ **Simple or One-way table :** When the data are tabulated to one characteristics. For example, tabulation of data on population of world classified by one characteristic like religion.

❖ **Double or Two-way table :** When the data are tabulated according to two characteristics. For example, tabulation of data on population of world classified by two characteristic like religion and sex.

❖ **Complex or multi-way table :** When the data are tabulated according to many characteristics. For example, tabulation of data on population of world classified by two characteristic like religion, sex, and literacy.

Construction of Statistical Table

- A statistical table has at least four major parts and some other minor parts.

- The Title
- The Box Head (column captions)
- The Stub (row captions)
- The Body
- Prefatory Notes
- Foot Notes
- Source Notes

General Sketch Table

THE TITLE
(Prefatory Notes)

	Box Head		
Row Caption	**Column Caption**		
Stub Entries	**The Body**		

Foot Notes…
Source Notes…

General Sketch Table

THE TITLE

- A title is the main heading written in capital shown at the top of the table.
- It must explain the contents of the table and throw light on the table as whole.
- Different parts of the heading can be separated by commas and no full stop should be used in the little.

THE TITLE

(Prefatory Notes)

	Box Head		
Row Caption	**Column Caption**		
Stub Entries	**The Body**		

Foot Notes…
Source Notes…

General Sketch Table

THE Box Head
(Column Captions)

- The vertical heading and subheading of the column are called columns captions.
- The spaces where these column headings are written is called box head.
- Only the first letter of the box head is in capital letters and the remaining words must be written in small letters.

THE TITLE
(Prefatory Notes)

	Box Head		
Row Caption	**Column Caption**		
Stub Entries	**The Body**		

Foot Notes…
Source Notes…

General Sketch Table

THE Stub

(Row Captions)

- The horizontal headings and sub-heading of the row are called row captions.
- The space where these row headings are written is called stub.

THE TITLE

(Prefatory Notes)

	Box Head		
Row Caption	**Column Caption**		
Stub Entries	**The Body**		

Foot Notes…
Source Notes…

General Sketch Table

THE Body

- It is the main part of the table which contains the numerical information classified with respect to row and column captions.
- Prefatory note: A statement given below the title and enclosed in brackets usually describe the units of measurement.

THE TITLE

(Prefatory Notes)

	Box Head		
Row Caption	**Column Caption**		
Stub Entries	**The Body**		

Foot Notes…

Source Notes…

General Sketch Table

Foot Notes

- It appears immediately below the body of the table providing the further additional explanation.
- The source notes is given at the end of the table indicating the source from where the information has been taken.

THE TITLE

(Prefatory Notes)

	Box Head		
Row Caption	**Column Caption**		
Stub Entries	**The Body**		

Foot Notes…

Source Notes…

General Rules of Tabulation

- A table should be simple and attractive. A complex table may be broken into relatively simple tables.
- Headings for columns and rows should be proper and clear.
- Suitable approximation may be adopted, and figures may be rounded off. But this should be mentioned in the prefatory note or in the foot note.
- The unit of measurement and nature of data should be well defined.

Organizing Data via Frequency Tables

- One method for simplifying and organizing data is to construct a **frequency distribution.**
- **A frequency distribution** is the organization of raw data in table form, using classes or groups and frequencies.
- **Class Frequency:** The number of observations falling in a particular class is called frequency or Simple frequency, denoted by 'f'.
- **Grouped Data:** Data presented in the form of frequency distribution is called grouped data.
- **Why use frequency distribution?**
- To convert raw data into meaningful form. It allows for quick visual interpretation of the data. It is drawn for both qualitative and quantitative data

Why Use Frequency Distributions?

- A frequency distribution is a way to summarize data. It condenses the raw data into a more meaningful form.

Types of Frequency Distributions

- There are three types of frequency distribution

  - **Ungrouped frequency distributions**
  - **Categorical frequency distributions**
  - **Grouped frequency distributions**

Frequency Distribution of Discrete Data

- **Example:** Number of children in 20 families.

  2, 3, 1, 3, 2, 5, 4, 1, 4, 2, 3, 5, 2, 5, 2, 1, 3, 1, 2, 0

  construct un-grouped or discrete frequency distribution.

Interpretation: There is
1 family with no children.
4 families with 1 children.
6 families with 2 children.
4 families with 3 children.
2 families with 4 children.
3 families with 1 children.

No of children	Tally	No of Families (Frequency)f
0	|	1
1	||||	4
2	||||| |	6
3	||||	4
4	||	2
5	|||	3
Total		20

Categorical Frequency Distributions

- Nominal- or ordinal-level data that can be placed in categories is organized in categorical frequency distributions. Religion, Political parties, blood type etc.
- **Example:** Suppose 30 patients blood was taken to determine their blood type, and the data is as follows.

  A, B, B, AB, O, O,A,B,AB,B,B,B,O,A,O,A,O,O,O,AB,AB,A,O,B,A,A,O, B,AB,AB

  Raw data, we have to arrange this data in a meaningful manner.

Class	Tally	Frequency	Percent %=f/n*100
A	𝍸 ||	7	=7/30*100= 23.33%
B	𝍸 |||	8	=8/30*100= 26.66%
O	𝍸 ||||	9	=9/30*100= 30%
AB	𝍸 |	6	=7/30*100= 20%
		Total 30	100%

We conclude more people have O blood type

Grouped Frequency Distributions

- Sometimes, when the data is continuous or covers a wide range of values, it becomes very burdensome to make a list of all values as in that case the list will be too long.

- When the range of the data is large, then we use grouped frequency distribution. Grouped frequency distribution is the organization of raw data in table form, using classes and frequencies. The largest data value that can be included in a class is the upper class limit for that class; the smallest data value that can be included is the lower class limit.

- Class boundaries separated the class.

Grouped Frequency Distributions

- The following are marks obtained by 30 students of a class in a test:

39, 25, 5,33, 19, 21, 12, 48, 13, 21, 9, 1, 10, 8, 12, 17, 19, 17, 41, 40, 12, 46, 37, 17, 27, 30, 6, 2, 23, 19

Grouped Frequency Distribution

Political Party Affiliations: Professor X asked his introductory statistics students to state their political party affiliations as P1, P2, P3 and P4. The responses of the 30 students in a class are:

P2 P1 P4 P3 P1 P4 P1 P2 P3 P1
P3 P1 P3 P2 P1 P4 P1 P3 P4 P3
P2 P3 P1 P3 P4 P3 P1 P3 P4 P2

People Power Party (P1), Democratic Party of Korea (P2), Justice Party (P3), National revolutionary Party (P4)

Construct a frequency distribution.

Party	Tally	Freq (f)
P1	𝍸 | | | |	9
P2	𝍸	5
P3	𝍸 𝍸	10
P4	𝍸 |	6
Total		**30**

Interpretation:

Out of 30 students in the class,

9 are in favor of P1

5 are in favor of P2

10 are in favor of P3

6 are in favor of P4

Grouped Frequency Distribution

- **Relative Frequency** is the ratio of the frequency to the total number of observations.

Relative frequency = Frequency/Number of observations

Example:

Relative frequency of students who favored P1=9/30=0.3=30%

Relative frequency of students who favored P2=5/30=0.167=16.67%

Relative frequency of students who favored P3=10/30=0.333=33.33%

Relative frequency of students who favored P4=6/30=0.2=20%

Grouped Frequency Distribution

Party Affiliation Example:

Interpretation: Out of 30 students in the class,

30% are in favor of P1

16.7% are in favor of P2

33.3% are in favor of P3

20% are in favor of P4.

Party	Freq (f)	Relative Freq
P3	10	10/30=0.3333
P1	9	9/30=0.30
P4	6	6/30=0.20
P2	5	5/30=0.1667
Total	**30**	**1**

Grouped Frequency Distribution

- **Example:** Ages of the 54 wealthiest people in the world. (Just assume, the data is not real)

54,34,45,56,67,81,45,23,34,56,76,54,32,82,84,34,56,45,76,34,24,23,85,27,76,53,44,37,82,56,78,57,41,83,48,46,58,59,63,64,76,58,42,41,35,65,45,78,70,60,50,40,35,50

Class Limit	Tally	Frequency(f)
23-29	||||	4
30-36	𝍸 ||	7
37-43	||||	4
44-50	𝍸 𝍸	10
51-57	𝍸 |||	8
58-64	𝍸 |	6
65-71	|||	3
72-78	𝍸 |	6
79-85	𝍸 |	6
		54

Basic rules for classes

- There should be between 5 and 20 classes. (no hard and fast rule)
- The classes must be mutually exclusive. This means that no data value can fall into two different classes.
- The classes must be all inclusive or exhaustive. This means that all data values must be included.
- The classes must be continuous. There are no gaps in a frequency distribution. Classes that have no values in them must be included (unless it's the first or last class which are dropped).

Class Limit
23-29
30-36
37-43
44-50
51-57
58-64
65-71
72-78
79-85

Creating a Grouped Frequency Distributions

- Find the largest and smallest values.
- Compute the range=maximum-minimum
- Select the number of classes desired. This is usually between 5 and 20.
- Find the class width by dividing the range by the number of classes and rounding up.
- To find the upper limit of the first class, subtract one from the lower limit of the second class. Then continue to add the class width to this upper limit to find the rest of the upper limits.
- Find the boundaries by subtracting 0.5 units from the lower limits and adding 0.5 units from the upper limits. The boundaries are also half-way between the upper limit of one class and the lower limit of the next class. Depending on what you're trying to accomplish, it may not be necessary to find the boundaries.
- Tally the data.
- Find the frequencies.
- Find the cumulative frequencies. Depending on what you're trying to accomplish, it may not be necessary to find the cumulative frequencies.
- If necessary, find the relative frequencies and/or relative cumulative frequencies.

Grouped Frequency Distribution

- **Example:** Ages of the 54 wealthiest people in the world. (Just assume, the data is not real)

54,34,45,56,67,81,45,23,34,56,76,54,32,82,84,34,56,45,76,34,24,23,85,27,76,53,44,37,82,56,78,57,41,83,48,46,58,59,63,64,76,58,42,41,35,65,45,78,70,60,50,40,35,50

Range=85-23=62/9=6.8 approx. 7 width

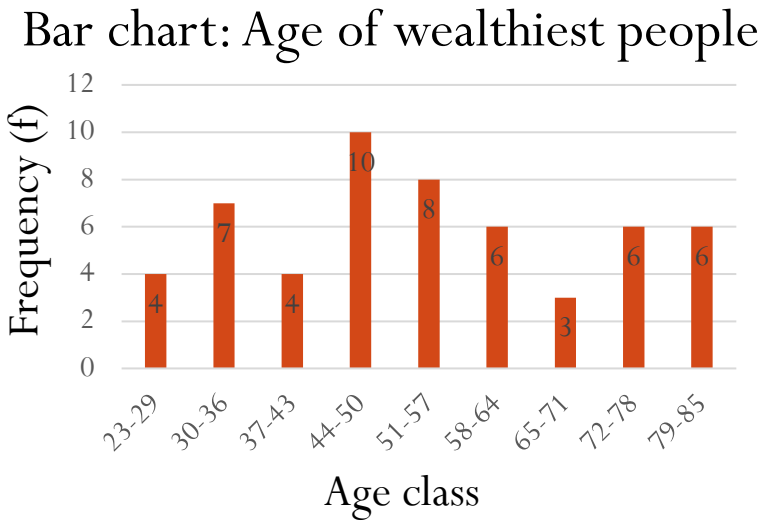
Class Limit	Class boundaries	Tally	Frequency(f)	Cumulative Frequency	Relative Frequency
23-29	22.5-29.5	||||	4	4	=4/54*100=>7.4%
30-36	29.5-36.5	𝍸 ||	7	11	=7/54*100=>12.96%
37-43	36.5-43.5	||||	4	15	=4/54*100=>7.4%
44-50	43.5-50.5	𝍸 𝍸	10	25	=10/54*100=>18.5%
51-57	50.5-57.5	𝍸 |||	8	33	=8/54*100=>14.8%
58-64	57.5-64.5	𝍸 |	6	39	=6/54*100=>11.1%
65-71	64.5-71.5	|||	3	42	=3/54*100=>5.5%
72-78	71.5-78.5	𝍸 |	6	48	=6/54*100=>11.1%
79-85	78.5-85.5	𝍸 |	6	54	=6/54*100=>11.1%
			54		=54/54*100=100%

Graphical Presentation of Data

- An old Saying "a picture is worth a thousand words".
- Graph or chart of a data set often provides the simplest and most efficient display.

- **Bar chart:** A Simple bar chart consists of horizontal or vertical bars of equal widths and lengths proportional to the values they represent.
- It displays graphically the same information concerning qualitative data that a frequency distribution shows in tabular form.

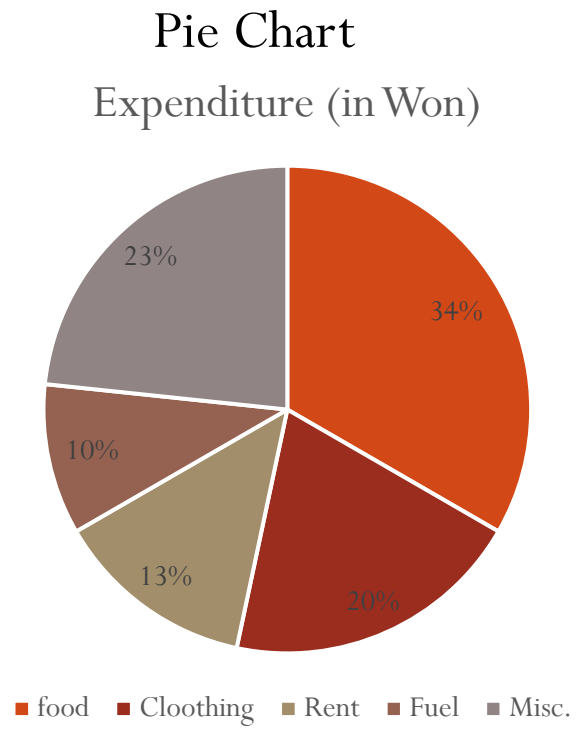
Class Limit	Frequency(f)
23-29	4
30-36	7
37-43	4
44-50	10
51-57	8
58-64	6
65-71	3
72-78	6
79-85	6



Graphical Presentation of Data

- **Pie Charts:** A Pie-Chart (also called sector diagram), is a graph consisting of a circle divided into sectors whose areas are proportional to the various parts into which whole quantity is divided.
- **Example:** Represent the expenditures on various items of a family by a pie chart.

Items	Expenditure (in 100000 won)
food	50
Clothing	30
Rent	20
Fuel	15
Misc.	35
Total	150



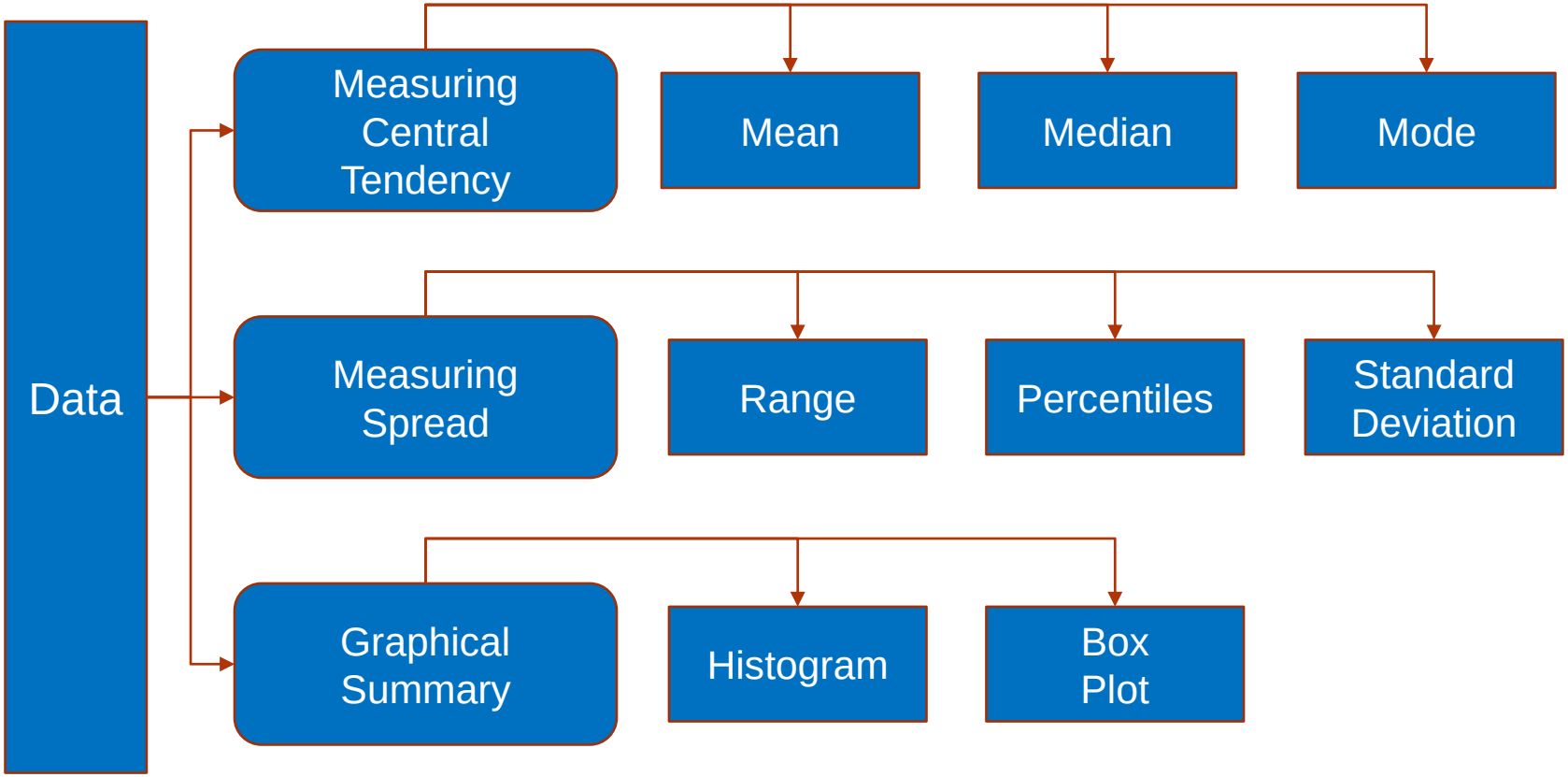
Measure of Central Tendency

Learning Objectives:

- Summarize data, using measures of central tendency, such as the
- Mean,
- Median,
- Mode,
- Midrange, and
- Weighted Mean

- Measures of average are called measures of central tendency.
- Examples: The average height of women is 5 feet and 3 inches.
  The average marks are 80%
  The average speed in the school zone is 30Km/hr etc.
- A statistic is a characteristic or measure obtained by using the data values from a sample.
- A parameter is a characteristic or measure obtained by using all the data values for a specific population.

Summarizing Data



Mean

- The mean is the quotient of the sum of the values and the total number of values.

- The Symbol $\bar{X}$ is used for sample mean. (Roman letters are used for statistics)

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_n}{n} = \frac{\sum X}{n}$$

- For a population, the Greek letter μ(mu) is used for the mean.

- Greek letters are used for parameters

$$\mu = \frac{X_1 + X_2 + X_3 + X_4 + X_n}{N} = \frac{\sum X}{N}$$

Mean

Example: The phone calls that a fire department responded to for a sample of 7 months is shown. Find the mean of the following data.

500, 650, 750, 450, 330,810, 655

Solution: we know that

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_n}{n} = \frac{\sum X}{n}$$

$$\bar{X} = \frac{500+650+750+450+330+810+655}{7} = \frac{4145}{7}$$

The mean is 592.14

The mean, in most cases, is not an actual data value

Mean

➢ **Example:**

A sample of the heights (in cm) of 5 students:
(170.90, 171.45, 168.68, 172.05, 166.27)

Solution: The mean of the sample is :

(170.90+ 171.45 + 168.68 + 172.05 +166.27) / 5 = 169.87

Example:

A simple random sample of five men is chosen from a large population of men, and their heights are measured. The five heights (in inches) are 65.51, 72.30, 68.31, 67.05, and 70.68. Find the sample mean.

Solution: The mean of the sample is :

$$\overline{X} = \frac{1}{5}(65.51 + 72.30 + 68.31 + 67.05 + 70.68) = 68.77 \text{ in.}$$

Find the Mean for Grouped Data

1. The frequency table shows the number of pets of 15 students. Find the mean number of pets.
Give your answer to 3 significant figures.

Number of pets	Frequency
0	2
1	8
2	4
3	1

2. The frequency table shows the price of a packet of biscuits in 16 different shops. Find the mean price.
Give your answer to 3 significant figures.

Price (p)	Frequency
59	10
60	2
61	2
62	1
63	1

Find the Mean for Grouped Data

The frequency table shows the marks scored in a test by 20 students. Find the mean.

Marks Scored	Frequency		
0 – 9	3		
10 – 19	5		
20 – 29	8		
30 – 39	4		

Find the Mean for Grouped Data

- **Below is a frequency distribution of books read per week. Find the mean.**

Class boundaries A	Frequency B f	Midpoint C X_m	D $X_m \cdot f$
3.5 – 8.5	5	6	30
8.5 – 13.5	6	11	66
13.5 – 18.5	4	16	64
18.5 – 23.5	3	21	63
23.5 – 28.5	5	26	130
28.5 – 33.5	4	31	124
	n=27		$\sum X_m \cdot f = 477$

$$\bar{X} = \frac{\sum X_m \cdot f}{n} = \frac{477}{27} = 17.66 \ \ books\ per\ week$$

Median

- The median is the midpoint of the data array. The symbol used for the median is MD.
- The raw data must be arranged in ascending order, and when the data is arranged, it is called data array.
- The median will be a specific value in the data set, if the data is in odd numbers, or will fall between two values if the data set is even.
- Arrange the data value in ascending order or increasing order.
- Determine the number of values in the data set.
- If numbers are odd, select the middle data value as the median.
- If numbers are even, find the mean of the two middle values. (add them and divide by 2)

Median

- Find the median of the following dataset.

A) 12, 34, 23, 45, 67 odd dataset

- Arrange in order

12, 23, 34, 45, 67 data array

B) 23, 56, 87, 34, 12, 76 even dataset

- Arrange in order

12, 23, 34, 56, 76, 87 data array

- Add 34+56

= 90/2

MD=45 is the median

Median From the Frequency Table

- In a frequency table, observations are already arranged in an ascending order

Median from the Frequency Table

Number of observations (n) is odd.

The median is the middle value, which is at position

$$\left(\frac{n+1}{2}\right)$$

Number of observations (n) is even.

The median is the average of the two middle values.

1. Find the value at position $\left(\frac{n}{2}\right)$
2. Find the value at position $\left(\frac{n}{2}\right)+1$
3. Find the average of the two values to get the median.

Median From the Frequency Table

How To Find The Median Of A Frequency Table When The Number Of Observations Is Odd?

Case 1: when the number of observations (n) is odd, then the median is the value at the $\left(\frac{n+1}{2}\right)^{th}$ position.

Example: The following is a frequency table of the score obtained in a mathematics quiz. Find the median score.

Score	0	1	2	3	4
Frequency	3	4	7	6	3

Solution:

Number of scores = 3 + 4 + 7 + 6 + 3 = 23 (odd number)

Since the number of scores is odd, the median is at the $\left(\frac{n+1}{2}\right)^{th} = \left(\frac{23+1}{2}\right)^{th} =$ 12^{th} position.

Median From the Frequency Table

How To Find The Median Of A Frequency Table When The Number Of Observations Is Odd?

Score	0	1	2	3	4
Frequency	3	4	7	6	3

To find out the 12^{th} position, we need to add up the frequencies as shown:

Score	0	1	2	3	4
Frequency	3	4	7	6	3
Position	3	3 + 4 = 7	7 + 7 =14		

The 12th position is after the 7th position but before the 14th position. So, the median is 2.

Median From the Frequency Table

How To Find The Median Of A Frequency Table When The Number Of Observations Is even?

Case 1: when the number of observations (n) is even, then the median is the value at the $\left(\frac{n}{2}\right)^{th}$ *and* $\left(\frac{n}{2}+1\right)^{th}$ position.

Example: The table is a frequency table of the scores obtained in a competition. Find the median score.

Scores	0	1	2	3	4
Frequency	11	9	5	10	15

Solution:

Number of scores = 11 + 9 + 5 + 10 + 15 = 50 (even number)

Since the number of scores is even, the median is at the average of the $\left(\frac{n}{2}\right)^{th} = \left(\frac{50}{2}\right)^{th} = 25^{th}$ position and $\left(\frac{n}{2}+1\right)^{th} = 26^{th}$ position.

Median From the Frequency Table

How To Find The Median Of A Frequency Table When The Number Of Observations Is even?

Scores	0	1	2	3	4
Frequency	11	9	5	10	15

To find out the 25^{th} position and 26^{th} position, we add up the frequencies as shown:

Scores	0	1	2	3	4
Frequency	11	9	5	10	15
Position	11	11 + 9 = 20	20 + 5 = 25	25 + 10 = 35	36 to 50

The score at the 25th position is 2 and the score at the 26th position is 3.

The median is the average of the scores at 25^{th} and 26^{th} positions= $\frac{2+3}{2} = 2.5$

Median of Grouped Frequency Table

The formula for the median of grouped data is:

$$M = l + \frac{(\frac{n}{2} - cf)}{f} \text{ x c}$$

Where

L = lower limit of the median class

n = total number of observations

cf = cumulative frequency of the class preceding the median class

f = frequency of the median class / each class

C= class length / class size

Median of Grouped Frequency Table

Example: Calculate the median for the following data:

Marks	0 - 20	20 - 40	40 - 60	60 - 80	80 - 100
Number of students	6	20	37	10	7

Solution: We need to calculate the cumulative frequencies to find the median.

Marks	**Number of students**	**Cumulative frequency**
0 - 20	6	6
20 - 40	20	26
40 - 60	37	63
60 - 80	10	73
80 - 100	7	80

Median of Grouped Frequency Table

- n= sum of cf = 80, N/2 = 80/2 = 40
- The cumulative frequency greater than 40 is 63 and the class is 40 - 60. Hence, the median class is 40 - 60.

$$M = l + \frac{(\frac{n}{2} - cf)}{f} \text{ x c}$$

$$M = 40 + \frac{(\frac{80}{2} - 26)}{37} \text{ x } 20$$

L = lower limit of the median class = 40

$$M = 40 + \frac{(40 - 26)}{37} \text{ x } 20$$

n = total number of observations= 80

cf = cumulative frequency of the class preceding the median class =26

$$M = 40 + \frac{(14)}{37} \text{ x } 20$$

f = frequency of the median class / each class = 37

$$M = 40 + \frac{(14)}{37} \text{ x } 20$$

C= class length / class size =20

$$M = 47.5$$

Therefore, the median is 47.5

Mode

- The number which appears most often in a dataset of numbers is called mode
- It is sometimes said to be the most important item.
- There may be no mode, one mode (unimodal), two modes (bimodal), or many modes (multimodal).
- **Example:**
- Find the mode from the following data set.

10, 23, 45, 67, 10, 10, 67

- Arrange in order but necessary for mode.

10, 10, 10, 23, 45, 67, 67

- Mode is 10 the dataset is said to be unimodal

Mode

- **Example:** Find the mode from the following data set.
  20,23,45,69,20,20,69,69

- Arrange in order
  20,20,20, 23,45,69,69,69
- Mode 20, and 69 the dataset is said to be bimodal

- **Example:** Find the mode from the following data set.
  24,61,30,69,20,24,69,61,45,48,45,24,45,61

- Arrange in order
  20,24,24,24,30,45,45,45,48,61,61,61,69,69
- Mode 24, 45 and 61 the dataset is said to be multimodal

Having more than two modes is called "Multimodel"

Mode

- **Example:** Find mode from frequency table (zero modes)

The following frequency table shows the number of pets owned by 10 different families in a certain neighborhood:

Pets	Frequency
0	2
1	2
2	2
3	2
4	2

- **Example:** Find mode from frequency table (1 mode)

The following frequency table shows the number of wins for soccer teams in certain league:

Wins	Frequency
0	2
1	3
2	4
3	1
4	2
5	3
6	2

Wins	Frequency
0	2
1	3
2	4
3	1
4	2
5	3
6	2

*The number of wins with the highest frequency is **2** wins.*

*Thus, the mode for this frequency table is **2**.*

Mode of grouped data

- **Formula**

$$Mode = l + \left[\frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}\right] h$$

where

l = lower limit of the modal class

f_m = frequency of the modal class

f_1 = frequency of class preceding the modal class

f_2 = frequency of class succeeding the modal class

h = width of the modal class

Mode of grouped data

- **Example:** The heights, in cm, of 50 students are recorded

Height (in cm)	125-130	130-135	135-140	140-145	145-150
Number of students	7	14	10	10	9

Here, the maximum frequency is 14 and the corresponding class is 130-135. So, 130-135 is the modal class such that

$$Mode = l + \left[\frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)}\right] h$$

l = lower limit of the modal class= 130
f_m = frequency of the modal class = 14
f_1 = frequency of class preceding the modal class= 7
f_2 = frequency of class succeeding the modal class= 10
h = width of the modal class= 5

Mode of grouped data

- **Example:** The heights, in cm, of 50 students are recorded

l = lower limit of the modal class= 130
f_m = frequency of the modal class = 14
f_1 = frequency of class preceding the modal class= 7
f_2 = frequency of class succeeding the modal class= 10
h = width of the modal class= 5

$$Mode = 130 + \left[\frac{14-7}{(14-7)+(14-10)}\right] * 5$$

$$Mode = 130 + \left[\frac{7}{7+4}\right] * 5$$

$$Mode = 133.18$$

Hence, the modal height = 133.18 cm.

Summary

- We discussed tabulation
- We discussed types of frequency distribution
- We discussed measure of central tendency
- We learned about sample statistics.
  - Mean
  - Median
  - Mode


Thank
You !