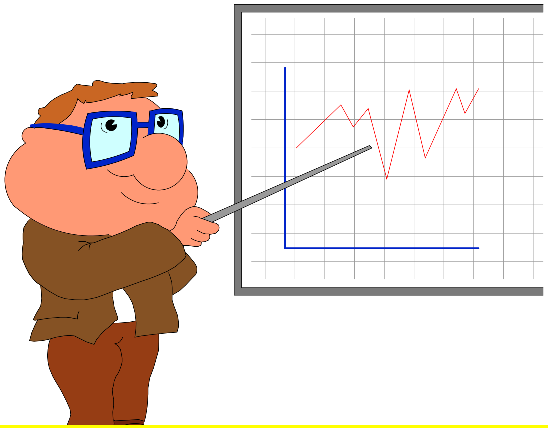
TAIBAH UNIVERSITY
Faculty of Science
Department of Math.

جامعة طيبة
كلية العلوم
قسم الرياضيات

Probability and Statistics for Engineers

STAT 301



Teacher :

Lesson

5

Measures of Location (Central Tendency)

Measures of Central Tendency

- **The data (observations) often tend to be concentrated around the center of the data.**
- **Some measures of location are: the mean, median and mode.**
- **These measures are considered as representatives (or typical values) of the data.**

Measures of Central Tendency

➢ **They are designed to give some quantitative measures of where the center of the data is in the sample.**

The Sample Mean:

- Is the most common measure of central tendency
- The sum of the values (positive , negative or zero) divided by the number of values
- Is called the Mean , Sample Mean , Arithmetic Mean and average.

The Sample Mean:

- **If the list is a statistical population, then the mean of that population is called a population mean ,denoted by μ.**
- **If the list is a statistical sample, we call the resulting statistic a sample mean. denoted by $\overline{X}$.**

The Sample Mean:

If $X_1, X_2, \ldots\ldots\ldots\ldots, X_n$ are the sample values, then the sample mean is:

$$\overline{X} = \frac{X_1 + X_2 + \ldots\ldots\ldots + X_n}{n}$$

Using summation:

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

The Sample Mean (Example 1):

Suppose that the following sample represents the ages (in year) of a sample of 3 men: $X_1 = 30,\ X_2 = 35,\ X_3 = 27.$

Then, the sample mean is:

$$\overline{X} = \frac{30+35+27}{3} = \frac{92}{3} = 30.67$$

The Sample Mean (Example 2):

- For what value of X will 8 and X have the same sample mean as 27 and 5?

Solution:

First, find the mean of 27 and 5:

$$\frac{27+5}{2}=16$$

Now, find the X value, knowing that the sample mean of X and 8 must be 16 :

$$\frac{X+8}{2}=16$$

cross multiply and solve: $32 = X + 8$ ⇨ $X = 24$

The Sample Mean (Example 3):

- On his first 5 Stat. tests, Omer received the following marks : 72, 86, 92, 63, and 77. What test mark must Omer earn on his sixth test so that his average for all six tests will be 80? .
- Solution

Set up an equation to represent the situation.

$$\frac{72+86+92+63+77+X}{6}=80 \Rightarrow X= 90$$

Omer must get a 90 on the sixth test.

The Sample Mean (Advantages):

- **Most popular measure in fields such as business, engineering and computer science.**
- **It is unique - there is only one answer.**
- **Useful when comparing sets of data.**

The Sample Mean (Disadvantages):

- Affected by extreme values (outliers)

Example.

- The sample mean of 2,3,4 is 3.
- The sample mean of 2,3,40 is 15.
- The mean increased from 3 to 15

because 40 is an extreme value.

The Sample Mean (Properties):

- **The sum of the deviations of the observations from their mean is always Zero**

$$\sum_{i=1}^{n}\left(X_i - \overline{X}\right) = 0$$

Example: Find the sum of the deviations of the values 3, 4, 6, 8,14 from their mean

$$\overline{X} = \frac{3+4+6+8+14}{5} = \frac{35}{5} = 7$$

$$\sum_{i=1}^{n}\left(X_i - \overline{X}\right) = (3-7) + (4-7) + (6-7) + (8-7) + (14-7) = 0$$

The Sample Median

- The purpose of the sample median is to reflect the central tendency of the sample in such a way that it is uninfluenced by extreme values or outliers

- The value which divides the data into two equal halves, with half of the data being lower than the median and half higher than the median.

The Sample Median (Steps):

If $X_1, X_2, \ldots\ldots\ldots\ldots, X_n$ are the sample values,

then the sample median computed as follows:

- **Sort the values into ascending order.**
- **If we have an odd number (n) of values,**

  the median is the middle value.

$$M_e = X_{\frac{1}{2}(n+1)}$$

The Sample Median (Steps):

- **If we have an even number of values, the median is the sample <u>mean</u> of the two middle values.**

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

The Sample Median (Example 1):

Compute The sample median of
(12, 24, 19, 20, 7) .

Solution

- **Sort the values into ascending order.**

  7,12, $\underline{19}$, 20, 24

- **Number of values is 5 is an odd the sample median is :**

$$M_e = X_{\frac{1}{2}(n+1)} = X_{\frac{1}{2}(5+1)} = X_3 = 19$$

The Sample Median (Example 2):

Compute The sample median of

(12, 24, 19, 20, 7 , 5) .

Solution

- **Sort the values into ascending order.**

5,7,<u>12</u>, <u>19</u>, 20, 24

- **Number of values is 6 is an even the sample median is :**

$$M_e = \frac{1}{2}\left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1}\right) = \frac{1}{2}\left(X_{\frac{6}{2}} + X_{\frac{6}{2}+1}\right)$$

$$= \frac{(X_3 + X_4)}{2} \qquad = \frac{(12+19)}{2} \qquad = 15.5$$

The Sample Median (Advantages):

- **Extreme values do not affect the median as strongly as they do to the mean.**
- **It is unique - there is only one answer.**
- **Useful when comparing sets of data.**

The Sample Median (Disadvantages):

Not as popular as sample mean

The Sample Mode :

- The value of a variable, which occurs with the highest frequency.
- The value of a distribution for which the frequency is maximum

The Sample Mode (Steps):

- Calculate the frequencies for all of the values in the data.
- The mode is the value (or values) with the highest frequency.

The Sample Mode (Examples):

- The sample mode of the list (1, 2, 2, 3, 3, 3, 4) is 3.
- The list (1, 2, 2, 3, 3, 5) has the two sample modes 2 and 3.
- No sample mode of the list (1, 6, 2, 7, 3, 5).

The Sample Mode (Advantages):

- **Extreme values do not affect the mode.**

Example:

- **The sample mode of the list (1, 2, 2, 3, 3, 3, 4) is 3.**
- **The sample mode of the list (1, 2, 2, 3, 3, 3, 4000) is 3.**

The Sample Mode (Disadvantages):

- Not as popular as mean and median.
- Not necessarily unique - may be more than one answer
- When no values repeat in the data set, the mode is every value and is useless.
- When there is more than one mode, it is difficult to interpret and/or compare.

Quartiles

The three values that split a set of ranked data values for a variable into four equal parts — quarters, or quartiles.

- The **First quartile (Q1)** is the value such that 25% of the ranked data are smaller and 75% are larger.

$$Q_1 = \begin{cases} x_{\frac{(n+1)}{4}} & \text{if } n \text{ is Odd} \\ \frac{1}{2}\left(x_{\frac{n}{4}} + x_{\frac{n}{4}+1}\right) & \text{if } n \text{ is Even} \end{cases}$$

- The **Second quartile (Q2)** is another name for the median.
- The **Third quartile (Q3)** is the value such that 75% of the ranked data are smaller and 25% are larger.

$$Q_3 = \begin{cases} x_{\frac{3(n+1)}{4}} & \text{if } n \text{ is Odd} \\ \frac{1}{2}\left(x_{\frac{3n}{4}} + x_{\frac{3n}{4}+1}\right) & \text{if } n \text{ is Even} \end{cases}$$

Example 4: Find the median, first quartile, third quartile of the following data of scores:
12 25 15 5 22 7 14 36 53 30 42

Solution: First, arrange the data in ascending order:

$$5 \quad 7 \quad \underset{\underset{Q_1}{\downarrow}}{12} \quad 14 \quad 15 \quad \underset{\underset{Median}{\downarrow}}{22} \quad 25 \quad 30 \quad \underset{\underset{Q_3}{\downarrow}}{36} \quad 42 \quad 53$$

$$Q_1 = x_{\frac{(n+1)}{4}} = x_{\frac{(11+1)}{4}} = x_3 = 12 \qquad Q_3 = x_{\frac{3(n+1)}{4}} = x_{\frac{3(11+1)}{4}} = x_{\frac{36}{4}} = x_9 = 36$$

Example 5: Find the median, first quartile, third quartile, if we add 65 to Example4.
Solution: First, arrange the data in ascending order:

$$5 \quad 7 \quad 12 \underset{\underset{Q_1}{\downarrow}}{} 14 \quad 15 \quad 22 \underset{\underset{Median}{\downarrow}}{} 25 \quad 30 \quad 36 \underset{\underset{Q_3}{\downarrow}}{} 42 \quad 53 \quad 65$$

$$Median(Q_2)=\frac{1}{2}\left(x_{\frac{2n}{4}}+x_{\frac{2n}{4}+1}\right)=\frac{1}{2}\left(x_{\frac{2*12}{4}}+x_{\frac{2*12}{4}+1}\right)=\frac{1}{2}(x_6+x_7)=\frac{1}{2}(22+25)=23.5$$

$$Q_1=\frac{1}{2}\left(x_{\frac{n}{4}}+x_{\frac{n}{4}+1}\right)=\frac{1}{2}\left(x_{\frac{12}{4}}+x_{\frac{12}{4}+1}\right)=\frac{1}{2}(x_3+x_4)=\frac{1}{2}(12+14)=13$$

$$Q_3=\frac{1}{2}\left(x_{\frac{3n}{4}}+x_{\frac{3n}{4}+1}\right)=\frac{1}{2}\left(x_{\frac{3*12}{4}}+x_{\frac{3*12}{4}+1}\right)=\frac{1}{2}(x_9+x_{10})=\frac{1}{2}(36+42)=39$$

When the result of this arithmetic is not a whole number (whether *n* is odd or even) use the following formulas for Q_1 and Q_3 :

$$Q_1=x_{\frac{(n+1)}{4}}$$

$$Q_3=x_{\frac{3(n+1)}{4}}$$

then select the ranked positions immediately below and above the number calculated. For example, for 10 values, the result (for the first quartile) would be 2.75 (10 + 1 is 11,11/4 is 2.75), and you would select the second and third ranked values. With these values, do the following:

1. Multiply the larger ranked value by the decimal fraction of the original result (0.75 in the example.)
2. Multiply the smaller ranked value by 1 minus the decimal fraction of the original result (0.25 for the example, because 1 – 0.75 is 0.25).
3. Add the two products to determine the quartile value.

Special case: Should the two ranked values selected be the same number, then the quartile is that number and you can skip the previous two multiplications and one addition.

Example 6: Find the first and third quartiles of the following data values:

Time	52	39	44	39	31	40	43	35	44	29

Solution: First, arrange the data in ascending order:

29 31 35 39 39 40 43 44 44 52

↓ ↓ ↓

Q_1 *Median* Q_3

$Q_1 = x_{\frac{(n+1)}{4}} = x_{\frac{(10+1)}{4}} = x_3 = 2.75$ The second and third ranked values are 31 and 35

1. Multiply the larger ranked value by the decimal fraction of the original result we get 35×0.75 =26.25
2. Multiply the smaller ranked value by 1 minus the decimal fraction of the original result we get 31×0.25 =7.75
3. First quartile is equal 26.25+7.75 =34.

$Q_1 = x_{\frac{(n+1)}{4}} = x_{\frac{(9+1)}{4}} = x_3 = 2.5$ The second and third ranked values are 35 and 39

4. Multiply the larger ranked value by the decimal fraction of the original result we get $39 \times 0.5 = 19.5$
5. Multiply the smaller ranked value by 1 minus the decimal fraction of the original result we get $35 \times 0.5 = 17.5$
6. First quartile is equal $19.5 + 17.5 = 37$.

$Q_3 = x_{\frac{3(n+1)}{4}} = x_{\frac{3(9+1)}{4}} = x_3 = 7.5$ The seventh and eighth ranked values are 44 and 44

1. Multiply the larger ranked value by the decimal fraction of the original result we get $44 \times 0.5 = 22$
2. Multiply the smaller ranked value by 1 minus the decimal fraction of the original result we get $44 \times 0.5 = 22$
3. Third quartile is equal $22 + 22 = 44$.

$$\frac{(n+1)}{4} = a.b \qquad Q_1 = .bX_{(a+1)} + (1-.b)X_{(a)}$$

$$\frac{3(n+1)}{4} = a.b \qquad Q_3 = .bX_{(a+1)} + (1-.b)X_{(a)}$$