Measures of Central Tendency (2)

Md. Ismail Hossain Riday

Descriptive measures

- In previous chapter, we discussed how a raw data set can be organized and summarized by tables and graphs.

- Another method of summarizing data set precisely is to compute number (a single number).

- Number that can be describe data sets are called descriptive measures

Combined Mean

- If $\bar{X}_1$ and $\bar{X}_2$ are the means with respective numbers of observations n_1 and n_2 of two data sets expressed in the same measuring units, then combined mean is given by,

$$\bar{X}_c = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2}$$

Combined Mean

- A consulting firm runs in two shifts. A random sample A of 13 employees has mean weekly salary 495$, and another random sample B of 10 employees has mean weekly salary 492$. Compute the arithmetic mean of the combined sample.

- Solution:

$$\bar{X}_c = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2} = 493.7\ \$$$

Weighted mean

$$\bar{X}_{WM} = \frac{\sum(w_i \times x_i)}{\sum w_i}$$

- The weighted mean is a special case of the arithmetic mean. It occurs when there are several observations of the same value.
- To explain- Suppose a burger company offers three different kinds of burger packages small, medium and large for Tk. 100, Tk. 125 and Tk. 150. Of the last 10 burgers sold 3 were small, 4 were medium and 3 were large. To find the mean price of the last 10 burger packages sold.

$$\bar{X}_{WM} = \frac{(3 \times 100) + (4 \times 150) + (3 \times 150)}{10} = 125$$

Self Practice

- Madina Construction Company pays its part time employees hourly basis. For different level of employee, the hourly rate is Tk 50, Tk 75 and Tk 90. There are 260 hourly employees, 140 of which are paid at Tk 50 rate, 100 at Tk 75 and 20 at the Tk 90 rate What is the mean hourly rate paid to the employees?

Median

- Middle value of the observation.
- After they have been arranged/ordered from smallest to largest.

Median for Ungrouped data

- There are two types of formula for ungrouped data

When "n" is odd:

$$Me = \left(\frac{n+1}{2}\right)^{th} Observation$$

When "n" is even:

$$Me = \frac{\left(\frac{n}{2}\right)^{th} Obs. + \left(\frac{n}{2}+1\right)^{th} Obs.}{2}$$

"n" is odd

- Step 1: Organize in ascending order
- Step 2: $Me = \left(\frac{n+1}{2}\right)^{th}\ Observation$
- For example: 5, 3, 9, 2, 7, 5, 8 are exam score of section "A".

Organize in ascending order:

2, 3, 5, 5, 7, 8, 9

Here, $n = 7$ is an odd number. Then the median can be written as,

$$Me = \left(\frac{n+1}{2}\right)^{th}\ Obs.$$

$$\therefore Me = 5$$

Comment: The median score of section "A" is 5

“n” is even

- Step 1: Organize in ascending order
- Step 2: $Me = \frac{\left(\frac{n}{2}\right)^{th} Obs. + \left(\frac{n}{2}+1\right)^{th} Obs.}{2}$
- For example: 5, 3, 9, 2, 7, 5 are exam score of section “A”.

Organize in ascending order:

2, 3, 5, 7, 8, 9

Here, $n = 6$ is an even number. Then the median can be written as,

$Me = 6$

Comment: The median score of section “A” is 6

Median for grouped data

$$Me = L_m + \frac{\frac{N}{2} - F_c}{f_m} \times c$$

$L_m = Lower\ limit\ of\ median\ class$

$N = Total\ number\ of\ observations$

$F_c = Cumultative\ frequency\ of\ pre - median\ class$

$f_m = Frequency\ of\ median\ class$

$c = Class\ interval$

Steps…

- Prepare a less than type cumulative frequency distribution.
- Determine $\frac{N}{2}$, where N is the total frequency.
- Locate the median class whose cumulative frequency includes the value of $\frac{N}{2}$.
- Determine the value of $L_m, F_c, f_m, and\ c$.

Median for grouped data

Class	Frequency	CF_i
5-9	4	4
9-13	3	7
13-17	3	10
17-21	3	13

Here, $N = 13$.

Now, $\frac{N}{2} = \frac{13}{2} = 6.5$

So, the median class is (9-13)

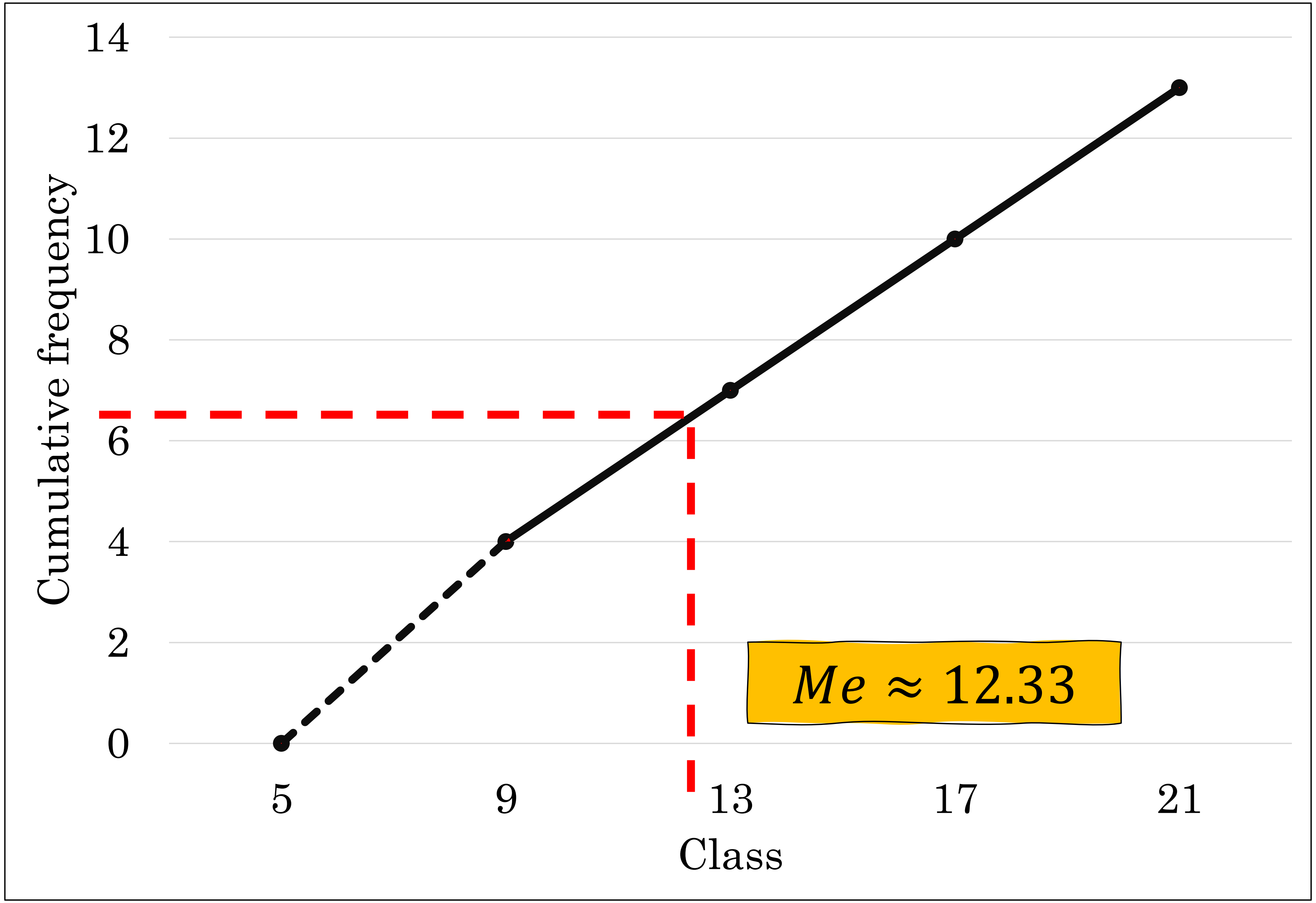
We know that,

$$Me = L_m + \frac{\frac{N}{2} - F_c}{f_m} \times c$$

$$\therefore Me = 9 + \frac{6.5 - 4}{3} \times 4 = 12.33$$

Median from graph

$$\frac{N}{2} = 6.5$$



Mode

- One mode = Unimodal
- Two mode = Bi-modal
- More than two = Multimodal

- Most frequent value
- Occurs more than one times.

- For example: 1, 2, 2, 5, 5, 3, 6, 6, 6, 4

- Here the most frequent value is 6.
- Thus the mode value is 6, and this is a unimodal data.

Mode

- Example: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

- There is no value which occur more than one time. Thus there is no "Mode" in this data set.

Mode for grouped data

$$Mo = L_o + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times c$$

L_o = Lower limit of modal class

Δ_1 = Difference between frequency of the modal class and pre-modal class.

Δ_2 = Difference between frequency of the modal class and post-modal class.

$c = Class\ interval$

Mode for grouped data

Class	f_i
10-20	5
20-30	8
30-40	12
40-50	7
50-60	9

Here, the class with highest frequency is (30-40). Thus, this class is our modal class.

Now, $Mo = L_o + \frac{\Delta_1}{\Delta_1+\Delta_2} \times c$

$\therefore Mo = 30 + \frac{4}{4+5} \times 10 = 34.44$

Some points...

- Does all data have a Mean, median, and mode?
- Every set of continuous data possesses a median, mode, and mean.
- When considering ordinal data, it encompasses solely a median and mode.
- Nominal data solely involves a mode

Measures of Location

- We have learned that median divided a set of data into two equal parts.

- In the same way, we can divides a set of data into

1. Four equal parts Quartiles
2. Ten equal parts Deciles
3. Hundred equal parts Percentiles

Q, D, and P

- Quartiles: Divide the data set into four equal parts. So, we get 3 quartile points: Q_1, Q_2, and Q_3

- Deciles: Divide the data set into ten equal parts. So, we get 9 decile points: D_1, D_2, D_3, … , D_9.

- Percentiles: Divide the data set into hundred equal parts. So, we get 99 percentile points: P_1, P_2, P_3, … , P_{99}

Q, D, and P (Ungrouped)

$$position\ of\ Q_i = \frac{i \times N}{4}$$

$i = 1, 2, 3$
$N = Total\ number\ of\ observations$

$$position\ of\ D_i = \frac{i \times N}{10}$$

$i = 1, 2, 3, 4, \dots, 9$
$N = Total\ number\ of\ observations$

$$position\ of\ P_i = \frac{i \times N}{100}$$

$i = 1, 2, 3, 4, \dots, 99$
$N = Total\ number\ of\ observations$

Steps...

$$position\ of\ D_i = \frac{i \times N}{10}$$

$$position\ of\ P_i = \frac{i \times N}{100}$$

- Arrange data set from smallest to largest
- Identify the position of Q_i by utilizing the formula,

$$J = \frac{(i \times n)}{4}$$

- If J is the integer value, then

$$\frac{\left(J^{th} Observation + (J+1)^{th} Observation\right)}{2}$$

- If J is not integer value, then take the next integer value as position.

Example (Ungrouped)

Data: 20, 22, 27, 33, 23

Organize the data into ascending order, 20, 22, 23, 27, 33

Now, position of $Q_1 = \frac{i \times N}{4} = 1.25$

Since, the position value is not integer, thus we should go for next integer value.

$$\therefore Q_1 = 2nd\ obs. = 22$$

Data: 20, 22, 27, 23

Organize the data into ascending order, 20, 22, 23, 27

Now, position of $Q_1 = \frac{i \times N}{4} = 1$

Since, the position value is integer

$$\therefore Q_1 = \frac{1st\ Obs. + 2nd\ Obs}{2} = 21$$

Q, D, and P (Grouped)

$$Q_i = L_Q + \frac{\frac{i \times N}{4} - F_{Qc}}{f_Q} \times c$$

$L_Q = Lower\ limit\ of\ quartile\ class$

$N = Total\ number\ of\ observations$

$F_c = Cumultative\ frequency\ of\ pre - quartile\ class$

$f_Q = Frequency\ of\ quartile\ class$

$c = Class\ interval$

Q, D, and P (Grouped)

$$D_i = L_D + \frac{\frac{i \times N}{10} - F_{Dc}}{f_D} \times c$$

$L_D = Lower\ limit\ of\ decile\ class$

$N = Total\ number\ of\ observations$

$F_c = Cumultative\ frequency\ of\ pre - decile\ class$

$f_D = Frequency\ of\ decile\ class$

$c = Class\ interval$

Q, D, and P (Grouped)

$$P_i = L_P + \frac{\frac{i \times N}{100} - F_{Pc}}{f_P} \times c$$

$L_P = Lower\ limit\ of\ percentile\ class$

$N = Total\ number\ of\ observations$

$F_c = Cumultative\ frequency\ of\ pre - percentile\ class$

$f_P = Frequency\ of\ percentile\ class$

$c = Class\ interval$

Self Practice

Determine $Q_1, Q_2, and\ Q_3$ with interpretation.

- A social organization publishes information on the TV-viewing habits of Bangladeshi garments workers in a report on Television. A sample of 100 garments workers yielded the monthly viewing times, in hours, displayed on the table,

Ans: 60.9, 68, and 79

TV-viewing times (in hours)	f_i
<50	8
50-60	14
60-70	35
70-80	20
80-90	15
≥90	8

Mathematical exercise

To access additional mathematical problems, please refer to the PDF lecture notes.

Thank You