Lecture 2 – Grouped Data Calculation

1. Mean, Median and Mode
2. First Quantile, third Quantile and Interquantile Range.

Mean – Grouped Data

Example: The following table gives the frequency distribution of the number of orders received each day during the past 50 days at the office of a mail-order company. Calculate the mean.

Number of order	f
10 – 12	4
13 – 15	12
16 – 18	20
19 – 21	14
	$n = 50$

Solution:

Number of order	f	x	fx
10 – 12	4	11	44
13 – 15	12	14	168
16 – 18	20	17	340
19 – 21	14	20	280
	$n = 50$		= 832

X is the midpoint of the class. It is adding the class limits and divide by 2.

$$\bar{x} = \frac{\sum fx}{n} = \frac{832}{50} = 16.64$$

Median and Interquartile Range – Grouped Data

Step 1: Construct the cumulative frequency distribution.
Step 2: Decide the class that contain the median.
Class Median is the first class with the value of cumulative frequency equal at least n/2.
Step 3: Find the median by using the following formula:

$$\text{Median} = L_m + \left(\frac{\frac{n}{2} - F}{f_m} \right) i$$

Where:

n = the **total frequency**

F = the **cumulative frequency *before*** class median

f_m = the **frequency** of the class median

i = the class width

L_m = the **lower boundary** of the class median

Example: Based on the grouped data below, find the median:

Time to travel to work	**Frequency**
1 – 10	**8**
11 – 20	**14**
21 – 30	**12**
31 – 40	**9**
41 – 50	**7**

Solution:

1st Step: Construct the cumulative frequency distribution

Time to travel to work	**Frequency**	**Cumulative Frequency**
1 – 10	**8**	**8**
11 – 20	**14**	**22**
21 – 30	**12**	**34**
31 – 40	**9**	**43**
41 – 50	**7**	**50**

$$\frac{n}{2} = \frac{50}{2} = 25 \longrightarrow$$ class median is the 3rd class

So, $F = 22$, $f_m = 12$, $L_m = 20.5$ and $i = 10$

Therefore,

$$\text{Median} = L_m + \left(\frac{\frac{n}{2} - F}{f_m} \right) i$$

$$= 21.5 + \left(\frac{25 - 22}{12} \right) 10$$

$$= 24$$

Thus, 25 persons take less than 24 minutes to travel to work and another 25 persons take more than 24 minutes to travel to work.

Quartiles

Using the same method of calculation as in the Median, we can get Q_1 and Q_3 equation as follows:

$$Q_1 = L_{Q_1} + \left(\frac{\frac{n}{4} - F}{f_{Q_1}} \right) i \qquad Q_3 = L_{Q_3} + \left(\frac{\frac{3n}{4} - F}{f_{Q_3}} \right) i$$

Example: Based on the grouped data below, find the Interquartile Range

Time to travel to work	**Frequency**
1 – 10	**8**
11 – 20	**14**
21 – 30	**12**
31 – 40	**9**
41 – 50	**7**

Solution:

1st Step: *Construct the cumulative frequency distribution*

Time to travel to work	***Frequency***	***Cumulative Frequency***
1 – 10	***8***	***8***
11 – 20	***14***	***22***
21 – 30	***12***	***34***
31 – 40	***9***	***43***
41 – 50	*7*	*50*

2nd Step: Determine the Q_1 and Q_3

$$\text{Class } Q_1 = \frac{n}{4} = \frac{50}{4} = 12.5$$

Class Q_1 is the 2nd class

Therefore,

$$Q_1 = L_{Q_1} + \left(\frac{\frac{n}{4} - F}{f_{Q_1}} \right) i$$

$$= 10.5 + \left(\frac{12.5 - 8}{14} \right) 10$$

$$= 13.7143$$

$$\text{Class } Q_3 = \frac{3n}{4} = \frac{3(50)}{4} = 37.5$$

$$Q_3 = L_{Q_3} + \left(\frac{\frac{n}{4} - F}{f_{Q_3}} \right) i$$

Class Q_3 is the 4th class

Therefore,

$$= 30.5 + \left(\frac{37.5 - 34}{9} \right) 10$$

$$= 34.3889$$

Interquartile Range

$$IQR = Q_3 - Q_1$$

$IQR = Q_3 - Q_1$

calculate the IQ

$IQR = Q_3 - Q_1 = 34.3889 - 13.7143 = 20.6746$

Mode – Grouped Data

Mode

•Mode is the value that has the highest frequency in a data set.
•For grouped data, class mode (or, modal class) is the class with the highest frequency.
•To find mode for grouped data, use the following formula:

$$\text{Mode} = L_{mo} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) i$$

Where:

i is the class width

Δ_1 is the difference between the frequency of class mode and the frequency of the class **after** the class mode

Δ_2 is the difference between the frequency of class mode and the frequency of the class **before** the class mode

L_{mo} is the **lower boundary** of class mode

Calculation of Grouped Data - Mode

Example: Based on the grouped data below, find the mode

Time to travel to work	**Frequency**
1 – 10	**8**
11 – 20	**14**
21 – 30	**12**
31 – 40	**9**
41 – 50	**7**

Solution:

Based on the table,

$L_{mo} = 10.5$, $\Delta_1 = (14 - 8) = 6$, $\Delta_2 = (14 - 12) = 2$ and

$i = 10$

$$\text{Mode} = 10.5 + \left(\frac{6}{6+2}\right)10 = 17.5$$

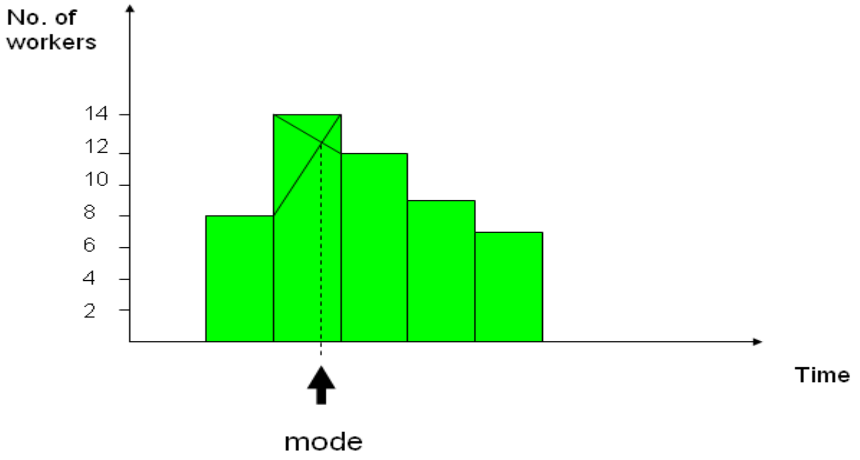
Mode can also be obtained from a histogram.

Step 1: Identify the modal class and the bar representing it

Step 2: Draw two cross lines as shown in the diagram.

Step 3: Drop a perpendicular from the intersection of the two lines until it touch the horizontal axis.

Step 4: Read the mode from the horizontal axis



Variance and Standard Deviation -Grouped Data

Population Variance:

$$\sigma^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{N}}{N}$$

Variance for sample data:

$$s^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{n}}{n-1}$$

Standard Deviation:

Population: $\sigma^2 = \sqrt{\sigma^2}$

Sample: $s^2 = \sqrt{s^2}$

Example: Find the variance and standard deviation for the following data:

No. of order	***f***
10 – 12	**4**
13 – 15	**12**
16 – 18	**20**
19 – 21	**14**
Total	**n = 50**

Solution:

No. of order	***f***	***x***	***fx***	***fx²***
10 – 12	**4**	**11**	**44**	**484**
13 – 15	**12**	**14**	**168**	2352
16 – 18	**20**	**17**	**340**	5780
19 – 21	**14**	**20**	**280**	5600
Total	**n = 50**		**832**	**14216**

Variance, $$s^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{n}}{n-1}$$

$$= \frac{14216 - \frac{(832)^2}{50}}{50-1}$$

$$= 7.5820$$

Standard Deviation, $$s = \sqrt{s^2} = \sqrt{7.5820} = 2.75$$

Thus, the standard deviation of the number of orders received at the office of this mail-order company during the past 50 days is 2.75.