Measures of Central Tendency, location and dispersion

1

Measures of Central Tendency

- Measured by averages
- Describe the point about which the various observed values cluster
- In mathematics, an **average**, or **central tendency** of a data set refers to a measure of the "middle" or "expected" value of the data set.

2

Measures of Central Tendency

- Arithmetic Mean
- Geometric Mean
- Weighted Mean
- Harmonic Mean
- Median
- Mode

3

Measures of Central Tendency

- Depends on whether observations are:
  - Grouped
  - Ungrouped

4

Grouped vs. Ungrouped Observations

- Grouped Observations
  - Observations are cast in a frequency distribution
- Ungrouped Observations
  - Observations are the exact values (raw data)

5

Arithmetic Mean of Ungrouped data

- sum of a set of observations, positive, negative or zero, divided by the number of observations.
- If we have "n" real numbers their arithmetic mean, denoted by $x_1, x_2, x_3, \ldots\ldots, x_n,$, can be expressed as:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots\ldots\ldots\ldots + x_n}{n}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

6

Arithmetic Mean for Ungrouped Data

$$\text{Mean} = \frac{\text{sum of the observations}}{\text{number of observations}}$$

7

Example 1

- A researcher is testing the extract X to 10 plants. The root lengths of the plants on that day were measured by the researcher and the observations were as follows:

Plant No.	Root Length in cms.
1	7.0
2	11.7
3	12.6
4	15.7
5	15.9
6	16.0
7	16.0
8	17.0
9	17.5
10	17.7
Total	147.1

8

Example 1

Mean = sum of observations / number of observations

Mean = 147.1 cm. / 10

Mean = 14.71 cm.

9

Arithmetic Mean of Grouped Data

- if $z_1, z_2, z_3, \ldots\ldots\ldots, z_k$ are the mid-values and $f_1, f_2, f_3, \ldots\ldots\ldots, f_k$ are the corresponding frequencies, where the subscript '*k*' stands for the number of classes, then the mean is

$$\bar{z} = \frac{\sum f_i z_i}{\sum f_i}$$

10

Example 2

- Determine the arithmetic mean on the distribution of root lengths (cm.) of 10 plants

Root lengths (cm.)	No. of plants (f)	Midpoint (X)
7 -	1	8.5
10 -	2	11.5
13 -	2	14.5
16 -	5	17.5
Total	10	

How was the midpoint determined?

11

Example 2

8.5 + 11.5 + 11.5 + 14.5 + 14.5 + 17.5 + 17.5 +

1 2 2

17.5 +17.5+ 17.5 = 146

5

OR (1 x 8.5)+(2 x 11.5)+(2 x 14.5)+(5 x 17.5)=146

12

Example 2

Mean = $\frac{\Sigma f x}{n}$

Mean = $\frac{146}{10}$

Mean = 14.6 cm.

13

Geometric Mean for Ungrouped Data

- Positive root of the product of observations. Symbolically,

$$G = (x_1 x_2 x_3 \cdots\cdots\cdots\cdots x_n)^{1/n}$$

- It is also often used for a set of numbers whose values are meant to be multiplied together or are exponential in nature, such as data on the growth of the human population or interest rates of a financial investment.

14

Example 3

- Find geometric mean of rate of growth: 34, 27, 45, 55, 22, 34

$$G = (x_1 x_2 x_3 \cdots\cdots\cdots\cdots x_n)^{1/n}$$

G = (34x27x45x55x22x34)$^{1/6}$

G =

15

Geometric Mean for Grouped Data

- If the "*n*" non-zero and positive variate-values occur $x_1, x_2, \ldots\ldots, x_n$ times, $f_1, f_2, \ldots\ldots, f_n$ respectively, then the geometric mean of the set of observations is defined by:

$$G = \left[x_1^{f_1}\ x_2^{f_2}\ \cdots\cdots\cdots x_n^{f_n}\right]^{\frac{1}{N}} = \left[\prod_{i=1}^{n} x_i^{f_i}\right]^{\frac{1}{N}}$$

Where $N = \sum_{i=1}^{n} f_i$

16

Geometric Mean

Ungroup Data

$$G = \sqrt{(x_1 x_2 x_3 \cdots\cdots\cdots\cdots x_n)}$$

$$G = AntiLog\left(\frac{1}{N}\sum_{i=1}^{n} Log\ x_i\right)$$

Group Data

$$G = \sqrt{(x_1^{f_1} x_2^{f_2} x_3^{f_3} \cdots\cdots\cdots\cdots x_n)}$$

$$G = AntiLog\left(\frac{1}{N}\sum_{i=1}^{n} f_i\ Log\ x_i\right)$$

17

Harmonic Mean

- Also called Subcontrary Mean
- One of the several kinds of averages
- Typically, it is appropriate for situations when the average of rates is desired.
- The harmonic mean is the number of variables divided by the sum of the reciprocals of the variables. Useful for ratios such as speed (=distance/time) etc.

18

Harmonic Mean

- The harmonic mean *H* of the positive real numbers $x_1, x_2, \ldots, x_n$ is defined to be

Ungroup Data	Group Data
$H = \dfrac{n}{\sum_{i=1}^{n} \dfrac{1}{x_i}}$	$H = \dfrac{n}{\sum_{i=1}^{n} \dfrac{f_i}{x_i}}$

19

Example 1

- Calculate the harmonic mean of the following numbers: 13.2, 14.2, 14.8, 15.2 and 16.1

20

Example 1

X	1/X
13.2	0.0758
14.2	0.0704
14.8	0.0676
15.2	0.0658
16.1	0.0621
Total	0.3417

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

H = 5/0.3417
H = 14.63

21

$\sum\left(\frac{f}{x}\right) = 1.4368$

Example 2

- Calculate the harmonic mean on the following:

Marks	X	f	f/X
30 – 39	34.5	2	0.0580
40 – 49	44.5	3	0.0674
50 -59	54.5	11	0.2018
60 – 69	64.5	20	0.3101
70 – 79	74.5	32	0.4295
80 – 89	84.5	25	0.2959
90 - 99	94.5	7	0.0741
Total		100	1.4368

22

Example 2

$$H = \frac{\sum f}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$$

H = 100/ 1.4368

H = 69.60

23

Weighted Mean

- The Weighted mean of the positive real numbers $x_1, x_2, ..., x_n$ with their weight $w_1, w_2, ..., w_n$ is defined to be

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

24

Example 1

- Situation:
- The College recently listed the names of people who had worked at the college for 15, 20, 25, and 30 years. There were 8 people who had worked for 15 years, 5 people for 20 years, 4 people for 25 years, and 1 person for 30 years. Determine the weighted mean length of service for these 18 people.

25

Example 1

- Solution:

If we wanted to find the mean length of service for these 18 people, we could add 15 + 15 +…+ 15 + 20 +…+ 20 + 25 +…+ 25 + 30, then divide the total by 18.

But we could also use the weighted mean. If a value is repeated, we multiply it by the number of times it appears in the list. We repeat this for all values in the list, add these products, then divide by the total number of values.

26

Example 1

- The number 15 appears 8 times in the above list. 15 is a value, and 8 is its weight. 20 is a value, and 5 is its weight. Here's the formula.

$$\bar{x}_w = \frac{\sum w \cdot x}{\sum w}$$

x is the individual value

w is the individual value's weight

27

Example 1

$$\begin{aligned}\bar{x}_w &= \frac{\sum w \cdot x}{\sum w} \\ &= \frac{8 \cdot 15 + 5 \cdot 20 + 4 \cdot 25 + 1 \cdot 30}{8+5+4+1} \\ &= \frac{120+100+100+30}{18} \\ &= \frac{350}{18} \\ &\approx 19.5 \text{ years of service}\end{aligned}$$

28

Median

- Middle value of the observations
- If n is odd, middlemost observation
- If n is even, mean of the two middlemost observations
- Steps:
  - Arrange observations in an array from lowest to highest
  - Identify the middlemost observation

29

Example 1

- Determine the median on the set of observations:

7.0 11.7 12.6 15.7 15.9 16.0 16.0 17.0 17.5 17.7

2 middlemost observations

$$\text{Median} = \frac{15.9 + 16.0}{2}$$

Median = 15.95

30

Example 2

- Determine the median on the set of observations:

7.0 11.7 12.6 15.7 15.9 16.0 16.0 17.0 17.5

5 lowest — 5 highest

Median = 15.9

31

Median for Grouped Data

$$M_e = L_o + \frac{h}{f_o}\left(\frac{n}{2} - F\right)$$

- L_0 = Lower class boundary of the median class
- h = Width of the median class
- f_0 = Frequency of the median class
- F = Cumulative frequency of the pre-median class

32

Median for Grouped Data

- Median class – first class whose cumulative frequency is greater than or equal to n/2.

Root lengths (cm)	No. of plants	Cumulative Frequency (CF)
7 -	1	1
10 -	2	3
13 -	2	5
16 -	5	10
Total	10	

Determine the Median?

33

Steps to find Median for Grouped Data

1. Compute the less than type cumulative frequencies.
2. Determine N/2 , one-half of the total number of cases.
3. Locate the median class for which the cumulative frequency is more than N/2 .
4. Determine the lower limit of the median class. This is L_0.
5. Sum the frequencies of all classes prior to the median class. This is *F*.
6. Determine the frequency of the median class. This is f_0.
7. Determine the class width of the median class. This is *h*.

34

Median for Grouped Data

- Median class = 10/2 = 5

Root lengths (cm)	No. of plants	Cumulative Frequency (CF)
7 -	1	1
10 -	2	3
13 -	2	5
16 -	5	10
Total	10	

$\geq n/2 = 5$

35

Median for Grouped Data

$$\text{Median} = Lo + \frac{h[\,n/2 - CF]}{fo}$$

$$\text{Median} = 13 + \frac{3[5 - 3]}{2}$$

Median = 16 cms.

36

Mode

- Value of a distribution for which the frequency is maximum.
- Value of a variable, which occurs with the highest frequency.
- So the mode of the list (1, 2, 2, 3, 3, 3, 4) is 3. The mode is not necessarily well defined. The list (1, 2, 2, 3, 3, 5) has the two modes 2 and 3.

37

Mode for Grouped Data

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} h$$

- L_1 = Lower boundary of modal class
- Δ_1 = difference of frequency between modal class and class before it
- Δ_2 = difference of frequency between modal class and class after
- H = class interval

38

Mode for Grouped Data

- Steps:
  - Find the modal class which has highest frequency
  - L_0 = Lower class boundary of modal class
  - h = Interval of modal class
  - Δ_1 = difference of frequency of modal class and class before modal class
  - Δ_2 = difference of frequency of modal class and class after modal class

39

Example 1

Class interval	6-10	11-15	16-20	21-25	26-30
Frequency	3	5	7	2	4

Solution:
-The 16-20 class has the most number, hence this is the modal class

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} h$$

40

Example 1

- Lower limit of modal class is 16
- Frequency of modal class is 7
- Frequency of class preceding modal class is 5
- Frequency of class succeeding modal class is 2
- Size of class is 5 (h)

  Mode = 16 + [(7-5)/(7 – 5)+(7- 2)] x 5

  Mode = 16 + [2/(2 + 5)x 5

  Mode = 16 + [2/7](5)

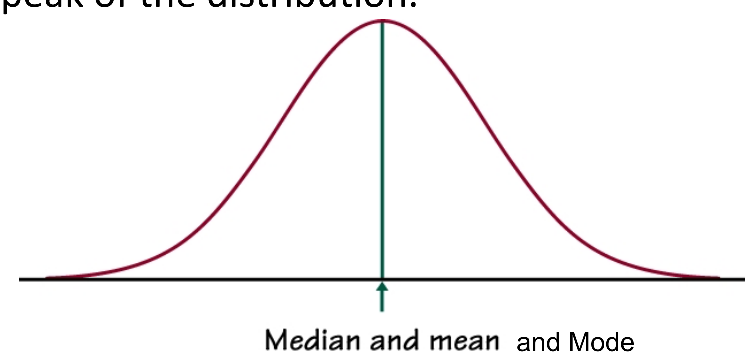
  Mode = 16 + 1.43

  Mode = 17.43

41

Characteristics of the Measures of Central Tendency

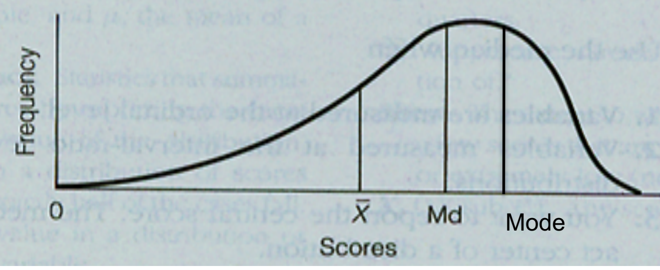
- When distribution is perfectly symmetrical, the mean, median and mode are all located at the center of the distribution which is also the peak of the distribution.



42

Characteristics of the Measures of Central Tendency

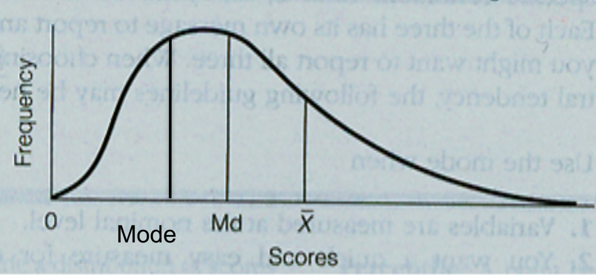
FIGURE 3.2 A NEGATIVELY SKEWED DISTRIBUTION (The mean is less than the median.)



43

Characteristics of the Measures of Central Tendency

FIGURE 3.1 A POSITIVELY SKEWED DISTRIBUTION (The mean is greater in value than the median.)



44

Choice of the Measures of Central Tendency

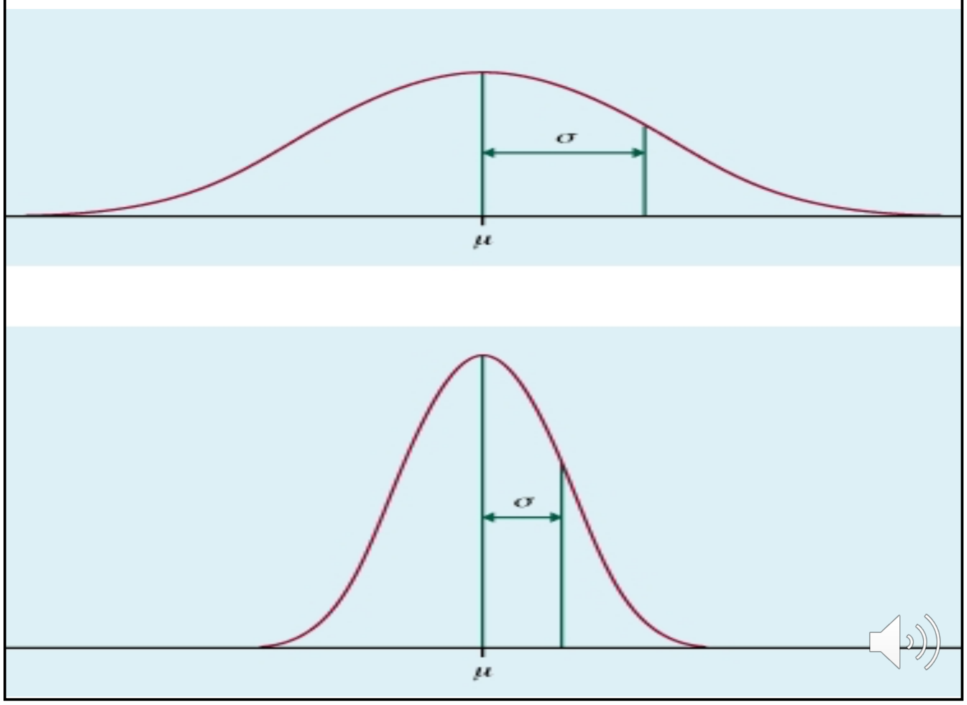
- Nature of the distribution
  - Symmetrical or skewed
- The concept of central tendency which is desired
  - Based on the objective of investigator

45

Measures of dispersion

- Tells us how the variable is spread
- Standard Deviations (σ) and Variance ($σ^2$)
- Range and coefficient of variation
- Statistical measures describing how different the data are from one another
- If the dispersion is small, it indicates high uniformity of the observations in the distribution.
- Absence of dispersion in the data indicates perfect uniformity. This situation arises when all observations in the distribution are identical.

46



47

Purpose of determining Dispersion

- Used to characterize a frequency distribution.
- Basis of comparison between two or more frequency distributions.
- The study of dispersion bears its importance from the fact that various distributions may have exactly the same averages, but substantial differences in their variability.

48

Range

- Simplest measure of variability
- Difference between highest and lowest value
- A big difference in range /bigger range
  - A high degree of variability
- A smaller difference in range/ smaller range
  - A small degree of variability

49

Range for Ungrouped Data

Range	=	Highest observation	–	lowest observation

50

Range for Grouped Data

Range = Upper boundary of the last class – Lower boundary of the first class

51

Percentile Range

- Difference between 10 to 90 percentile.
- It is established by excluding the highest and the lowest 10 percent of the items, and is the difference between the largest and the smallest values of the remaining 80 percent of the items.

$$P_{10}^{90} = P_{90} - P_{10}$$

52

Quartile Deviation

- A measure similar to the special range (Q) is the inter-quartile range . It is the difference between the third quartile (Q_3) and the first quartile (Q_1). Thus

$$Q = Q_3 - Q_1$$

- The inter-quartile range is frequently reduced to the measure of semi-interquartile range, known as the quartile deviation (QD), by dividing it by 2. Thus

$$QD = \frac{Q_3 - Q_1}{2}$$

53

Mean Deviation

The key concept for describing normal distributions and making predictions from them is called **deviation from the mean**.

We could just calculate the average distance between each observation and the mean.

- We must take the absolute value of the distance, otherwise they would just cancel out to zero!

Formula:

$$\sum \frac{|\overline{X} - X_i|}{n}$$

54

Mean Deviation: An Example

Data: X = {6, 10, 5, 4, 9, 8} $\bar{X} = 42 / 6 = 7$

$\bar{X} - X_i$	Abs. Dev.
7 – 6	1
7 – 10	3
7 – 5	2
7 – 4	3
7 – 9	2
7 – 8	1
Total:	12

1. Compute $\bar{X}$ (Average)
2. Compute $\bar{X} - X$ and take the Absolute Value to get Absolute Deviations
3. Sum the Absolute Deviations
4. Divide the sum of the absolute deviations by N

12 / 6 = 2

55

What Does it Mean?

- On Average, each observation is two units away from the mean.

Is it Really that Easy?

- No!
- Absolute values are difficult to manipulate algebraically
- Absolute values cause enormous problems for calculus (Discontinuity)
- We need something else…

56

Variance and Standard Deviation

- Instead of taking the absolute value, we square the deviations from the mean. This yields a positive value.
- This will result in measures we call the Variance and the Standard Deviation

Sample-	Population-
s: Standard Deviation	σ: Standard Deviation
s^2: Variance	σ^2: Variance

57

Standard Deviation

- Positive square root of the mean-square deviations of the observations from their arithmetic mean.

Population

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Sample

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N - 1}}$$

$$SD = \sqrt{\text{variance}}$$

58

Standard Deviation for Grouped Data

- SD is : $s = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{N}}$ Where $\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$
- Simplified formula

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

59

Example 1

- Determine Standard Deviation

Family No.	1	2	3	4	5	6	7	8	9	10
Size (x_i)	3	3	4	4	5	5	6	6	7	7

60

Example 1

Here, $\bar{x} = \frac{\sum x_i}{n} = \frac{50}{10} = 5$

Family No.	1	2	3	4	5	6	7	8	9	10	Total
x_i	3	3	4	4	5	5	6	6	7	7	50
$x_i - \bar{x}$	-2	-2	-1	-1	0	0	1	1	2	2	0
$(x_i - \bar{x})^2$	4	4	1	1	0	0	1	1	4	4	20
x_i^2	9	9	16	16	25	25	36	36	49	49	270

$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{20}{9} = 2.2,$ $s = \sqrt{2.2} = 1.48$

61

Example 2

- Determine Standard Deviation for Grouped Data

x_i	f_i	$f_i x_i$	$f_i(x_i)^2$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
3	2	6	18	-3	9	18
5	3	15	75	-1	1	3
7	2	14	98	1	1	2
8	2	16	128	2	4	8
9	1	9	81	3	9	9
Total	10	60	400	-	-	40

$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{60}{10} = 6$ $s^2 = \frac{\sum f(x_i - \bar{x})^2{}_i}{n-1} = \frac{40}{9} = 4.44$

62

Variance

- Measure of variability taking the mean as the reference point
- Takes into account the deviation of the individual observations from the mean
- Average of the squared deviations from the mean

63

Relative Measures of Dispersion

- To compare the extent of variation of different distributions whether having differing or identical units of measurements, it is necessary to consider some other measures that reduce the absolute deviation in some relative form.
- Expressed in the form of coefficients and are pure numbers, independent of the unit of measurements.

64

Relative Measures of Dispersion

- Coefficient of variation
- Coefficient of mean deviation
- Coefficient of range
- Coefficient of quartile deviation

65

Coefficient of Variation

- Expresses standard deviation as a percentage of the mean
- Computed as a ratio of the standard deviation of the distribution to the mean of the same distribution.

$$CV = \frac{s_x}{\bar{x}} \quad \text{X } 100$$

66

Example 3

	Height	weight
Mean	40 inch	10 kg
SD	5 inch	2 kg
CV	12.5%	20.0%

- Since the coefficient of variation for weight is greater than that of height, we would tend to conclude that weight has more variability than height in the population.

67

Coefficient of Mean Deviation

- As the mean deviation can be computed from mean, median, mode, or from any arbitrary value, a general formula for computing coefficient of mean deviation may be put as follows:

$$\text{Coefficient of mean deviation} = \frac{\text{Mean deviation}}{\textit{Mean}} \times 100$$

68

Coefficient of Range

- Relative measure corresponding to range and is obtained by the following formula:

$$\text{Coefficient of range} = \frac{L - S}{L + S} \times 100$$

- where, "L" and "S" are respectively the largest and the smallest observations in the data set.

69

Coefficient of Quartile Deviation

- Computed from the first and the third quartiles using the following formula:

$$\text{Coefficient of quartile deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

70

Measures of location

- Quartile, Deciles, Percentiles
  - Based on position of the series of location
  - Position of scores relative to the other scores as compared to the whole set of data

- Quartile
  - One fourth (1/4)
    - First (1/4), Second (1/2), Third (3/4)
- Decile
  - One tenth (1/10)
    - 10%, 20%, …90%
- Percentile
  - One of hundreds (1/100)
    - 1%, 2%, ….99%

71

Measures of Location

- Percentile – is one of the 99 values of a variable which divides the distribution into 100 equal parts.
- Decile – is one of the 9 values of a variable which divides the distribution into 10 equal parts.
- Quartile – is one of the 3 values of a variable which divides the distribution into 4 equal parts.

72

Quartile

$$Q_i = L_i + \frac{h}{f_i}\left(\frac{i}{4}n - F\right), i = 1, 2, 3$$

- L_0 = Lower limit of the *i*-th Quartile class
- *n* = Total number of observations in the distribution
- h = Class width of the *i*-th Quartile class
- f_i = Frequency of the *i*-th Quartile class
- F = Cumulative frequency of the class prior to the *i*-th quartile class

73

Decile

$$P_i = L_i + \frac{h}{f_i}\left(\frac{i}{10}n - F\right), i = 1, 2, 3$$

- L_0 = Lower limit of the *i*-th Decile class
- *n* = Total number of observations in the distribution
- h = Class width of the *i*-th Decile class
- f_i = Frequency of the *i*-th Decile class
- F = Cumulative frequency of the class prior to the *i*-th Decile class

74

Percentile

$$P_i = L_i + \frac{h}{f_i}\left(\frac{i}{100}n - F\right), i = 1, 2, 3$$

- L_0 = Lower limit of the *i*-th Percentile class
- *n* = Total number of observations in the distribution
- h = Class width of the *i*-th Percentile class
- f_i = Frequency of the *i*-th Percentile class
- F = Cumulative frequency of the class prior to the *i*-th Percentile class

75

Example 1

- Percentile of an Ungrouped Data

  Consider the following observations:

  11, 14, 17, 23, 27, 32, 40, 49, 54, 59, 71 and 80.

  Determine the 29th percentile?

76

Example 1

Observations: 11, 14, 17, 23, 27, 32, 40, 49, 54, 59, 71 and 80.

Determine the 29th percentile?

$$\frac{1}{100}(29 \times 12) = 3.48,$$

- Note that 3.48 is not an integer in the observations, thus the next higher integer 4 here will determine the 29th percentile value. On inspection $P_{29} = 23$

77

Example 2

- Find the 30th percentile in the grouped data

Age in years	Number of births	Cumulative number of births
14.5-	677	677
19.5-	1908	2585
24.5-	1737	4332
29.5-	1040	5362
34.5-	294	5656
39.5-	91	5747
44.5-	16	5763
All ages	5763	-

78

Example 2

- First determine the percentile class.
- If, N =5763, and we have to find 30th percentile, then percentile class will be the class which has cumulative frequency below:

$$\frac{i}{100}n = (30/100) \times 5763 = 1728.9.$$

79

Example 2

1728.9

- 30th Percentile class

Age in years	Number of births	Cumulative number of births
14.5-	677	677
19.5-	1908	2585
24.5-	1737	4332
29.5-	1040	5362
34.5-	294	5656
39.5-	91	5747
44.5-	16	5763
All ages	5763	-

80

Example 3

1st Quartile of ungrouped data **3rd Quartile of ungrouped data**

$$Q1 = \frac{n+1}{4}th \qquad Q3 = \frac{3(n+1)}{4}th$$

81

Example 3

- Given the n=9, and the observed values are:

  18 20 23 20 23 27 24 23 29

Arrange the values in ascending order of magnitude first

  18 20 20 23 23 23 24 27 29

The first quartile or the lower quartile is the

- (9+1)/4 = 2.5th term

2.5th term is (20 + 20)/2 = 20 OR 2.5 ~ 3rd term = 20

The 3rd quartile or the upper quartile is the

- 3(9+1)/4 =7.5th term

7.5th term is (24+27)/2 = 25.5 OR 7.5 ~ 8th term = 27

(Note: The median is the second quartile.)

82

Example 3

- Median or the 2^{nd} quartile

18 20 20 23 23 23 24 27 29

Median = (9+1)/2 = 5^{th} term

5^{th} term = 23

- Inter-quartile range:

= Upper quartile – Lower quartile

= 27 – 20

= 7

This means that the middle 50% of the data values range from 20 to 27

83

Example 4

- Given the n=8, and the observed values are:

2 4 5 7 7 8 10 12

Determine the first quartile and the 3^{rd} quartile:

*Arrange the observed values in ascending order of magnitude first

2 4 5 7 7 8 10 12

Q1 = (8+1)/4 = 2.25^{th} term

2.25^{th} term lies between 2^{nd} and 3^{rd} observation

2^{nd} observation = 4

3^{rd} observation = 5

Q1 = 4 + 0.25(5-4)

Q1 = 4 + 0.25

Q1 = 4.25

OR 2.25^{th} term ~ 2^{nd} term = 4

84

Example 4

- Determine Q3 in this set of observations:

  2 4 5 7 7 8 10 12

85

Example 5

- Decile Determination
  - Determine the 5th Decile on the following observations

  18 20 20 23 23 23 24 27 29 30

  Steps:

  *Arrange in ascending order of magnitude

  D = j (n+1)/10

  D = 5 (10+1)/10

  D = 55/10

  D = 5.5^{th} term

86

Example 5

- 5.5th term lies between the 5th and 6th term

  5th term = 23

  6th term = 23

  D = 23 + 0.5(23-23)

  D = 23

OR 5.5th term ~ 6th term = 23

87

Exercise 1

- Determine Decile, Percentile and Quartile for Grouped Data
- Find the 3rd quartile, 29th percentile and 1st decile in the following observations.

Marks	No. of students	Cumulative frequencies
40-	6	6
50-	11	17
60-	19	36
70-	17	53
80-	13	66
90-	4	70
Total	70	

88

Exercise 2

- Determine Decile, Percentile and Quartile for Ungrouped Data

Observations are:

5	5	4	3	7	9	10	11	13
10	8	12	14	16	18	20	3	4

89

Exercise 3

- Find the Arithmetic , Geometric and Harmonic Mean

Class	Frequency (f)	x	fx	f Log x	f / x
20-29	3	24.5	73.5	4.17	8.17
30-39	5	34.5	172.5	7.69	6.9
40-49	20	44.5	890	32.97	2.23
50-59	10	54.5	545	17.37	5.45
60-69	5	64.5	322.5	9.05	12.9
Sum	N=43		2003.5	71.24	35.64

90

Exercise 4

- Find Median

Age in years	Number of births	Cumulative number of births
14.5-19.5	677	677
19.5-24.5	1908	2585
24.5-29.5	1737	4332
29.5-34.5	1040	5362
34.5-39.5	294	5656
39.5-44.5	91	5747
44.5-49.5	16	5763
All ages	5763	-

91

Exercise 5

- Find Mean, Median and Mode of Ungroup Data

3 , 12 , 4 , 6 , 1 , 4 , 2 , 5 , 8

92

Exercise 6

- Find mode

Slope Angle (°)	Midpoint (x)	Frequency (f)	Midpoint x frequency (fx)
0-4	2	6	12
5-9	7	12	84
10-14	12	7	84
15-19	17	5	85
20-24	22	0	0
Total		n = 30	∑(fx) = 265

93

Exercise 7

- Determine the following:
  - Range, Percentile range, Quartile Range
  - Quartile deviation, Mean deviation, Standard deviation
  - Coefficient of variation, Coefficient of mean deviation, Coefficient of range, Coefficient of quartile deviation

94

Exercise 7

Marks	No. of students	Cumulative frequencies
40-	6	6
50-	11	17
60-	19	36
70-	17	53
80-	13	66
90-	4	70
Total	70	

95