

CHAPTER 6: Measures of Dispersion

- these are statistical measures that summarize the amount of spread or variation in the distribution of values in a variable.

- it describes how values are spread within the distribution

- it also describe how similar a set of scores are to each other.

- the *more similar* the scores are to each other, *the lower* the measures of dispersion will be.

- the *less similar* the scores are to each other, *the higher* the measures of dispersion will be.

- in general, the *more spread out* a distribution is, *the larger* the measure of dispersion will be.

RANGE

- it is the difference between the largest and smallest number in a set of observation.

- it is used mostly for quick and easy indication of variability.

- it can be used with ordinal or interval- ratio variables.

- the range is rarely used in scientific work as it is fairly insensitive.

- the range can be used when you are presenting your results to people with little or no knowledge of statistics.

- two diff sets of data may have same range.

1, 1, 1, 1, 9 vs 1, 3, 5, 7, 9

Range Formula:

Ungroup data

Range= Highest Score - Lowest Score

Group Data

Range= Highest Class Mark - Lowest Class Mark

INTER- QUARTILE RANGE (IQR)

- it is defined as the difference of the first and third quartile of a data set.

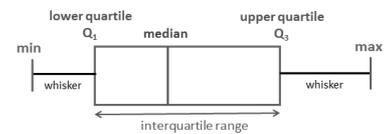
- it is a measure where the "middle fifty" lies in the data set.

- therefore, because it uses the middle 50%, it is not affected by outliers or extreme values.

INTER- QUARTILE RANGE (IQR)

Box and Whisker Plot

A box and whisker plot (also called a box plot) shows the five-number summary of a set of data: minimum, lower quartile, median, upper quartile, and maximum.



INTER- QUARTILE RANGE (IQR)

INTER-QUARTILE RANGE (IQR)

Formula:

Ungrouped /Grouped Data:

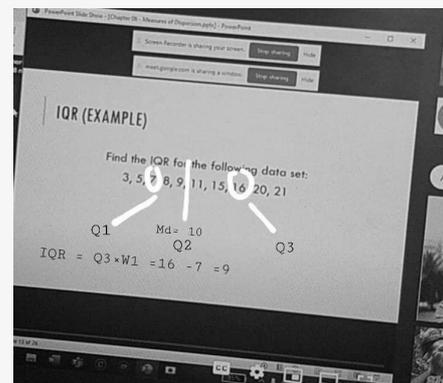
$$IQR = Q_3 - Q_1$$

INTER- QUARTILE RANGE (IQR)

Q3 = also known as UPPER QUARTILE

Q1= also known as LOWER QUARTILE

INTER- QUARTILE RANGE (IQR)



CHAPTER 6: Measures of Dispersion

2 IMPORTANT MEASURES OF DISPERSION

- variance
- standard deviation

VARIANCE

VARIANCE

- *It is defined as the average of the squared deviations.
- *It involves measuring the distance between each score and the mean.
- *The larger the variance is, the more the scores deviate, on average, away from the mean.
- *The smaller the variance is, the less the scores deviate, on average, from the mean.

VARIANCE

VARIANCE (FORMULA)

Ungrouped Data:

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{N}; \quad s^2 = \frac{\sum(x - \bar{x})^2}{N - 1}$$

↓

(population) (sample)

Grouped Data:

$$\sigma^2 = \frac{\sum f(x - \bar{x})^2}{N}; \quad s^2 = \frac{\sum f(x - \bar{x})^2}{N - 1}$$

↓

(population) (sample)

VARIANCE

VARIANCE

- *The variance formula tells us to subtract the mean from each score.
- *This difference is called **deviate** or **deviation**.
- *The deviate tells us how far a given score is from the typical, or average, score.

VARIANCE

VARIANCE (EXAMPLE)

The heights of the dogs (at the shoulders) are: 600mm, 470mm, 170mm, 430mm and 300mm. Find the variance.

VARIANCE

VARIANCE (SOLUTION)

Since we are only interested in the variance of the given dogs, we will use the population variance formula.

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{N}$$

$$\bar{x} = \frac{600 + 470 + 170 + 430 + 300}{5} = 394$$

$$\sigma^2 = \frac{(600 - 394)^2 + (470 - 394)^2 + (170 - 394)^2 + (430 - 394)^2 + (300 - 394)^2}{5}$$

$$\sigma^2 = 21704$$

STANDARD DEVIATION

CALCULATOR METHOD OF COMPUTING THE STD. DEVIATION

1. Key in MODE then 3 (DATA), then 1 (1-VARS).
2. Input data one by one by entering the score, then Δ . Continue until you have entered all the data.
3. Press Δ .
4. Press SRT then 1 (DATA) then 4 (VARS).
5. From the menu, choose the number corresponding to the sample standard deviation s_x .
6. Press Δ .

STANDARD DEVIATION

STANDARD DEVIATION (FORMULA)

Ungrouped Data:

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

Grouped Data:

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

STANDARD DEVIATION

STANDARD DEVIATION

- *It is the square root of the variance.
- *The most important and widely used measure of dispersion.
- *It should be used with interval-ratio variables but is often used with ordinal-level variables.
- *The lower the std. dev., the more consistent the data set is. Since it is the square root of the variance, the unit of the std. deviation is the same as the unit of the scores.
- *It can be used to determine how much of the data is near the mean.