Measures of Dispersion

Md. Ismail Hossain Riday

Dispersion

- In previous chapter, we learned how to summarize a raw data set using a single value that indicates the center of the distribution.

- It gives only partial information of the data set.

Dispersion

- For example, Record the examination results of two students
- Student A: 50, 49, 51, 50
- Student B: 100, 100, 0, 0

Dispersion

- For example, Record the examination results of two students
- Student A: 50, 49, 51, 50 $(\bar{X}_A = 50)$
- Student B: 100, 100, 0, 0

Dispersion

- For example, Record the examination results of two students
- Student A: 50, 49, 51, 50 $(\bar{X}_A = 50)$
- Student B: 100, 100, 0, 0 $(\bar{X}_B = 50)$

Dispersion

- For example, Record the examination results of two students

- Student A: 50, 49, 51, 50 $(\bar{X}_A = 50)$

- Student B: 100, 100, 0, 0 $(\bar{X}_B = 50)$

On average both students are equal in getting final exam score

So it is also important to describe "How much the observation vary from one to another

Dispersion/ Variation

Measures of Dispersion

- Dispersion may describe how much the observation vary from one to another.

- Descriptive measures that indicate the amount of variation in a data set are called measures of dispersion.

Why do we need?

- To determine the reliability of an average

- To compare the variability of two or more data sets.

Types of Measures of Dispersion

- Two types of measures of dispersion

1. Absolute measure

2. Relative measure

Absolute measures

- Only measure the inherent variation of a data set

Absolute measures

- Only measure the inherent variation of a data set

- Important absolute measures are-

1. Range
2. Mean Deviation
3. Variance
4. Quartile Deviation

Range

- $Range = Highest\ value - Lowest\ value$

Range

- $Range = Highest\ value - Lowest\ value$

- For example, weights of 5 newly born babies (in pounds)

$$7.5, 4.5, 10.1, 9.6, \mathit{and}\ 5.5$$

Range

- $Range = Highest\ value - Lowest\ value$

- For example, weights of 5 newly born babies (in pounds)

$$7.5, 4.5, 10.1, 9.6, and\ 5.5$$

$$(Range = ????)$$

Range

- $Range = Highest\ value - Lowest\ value$

- For example, weights of 5 newly born babies (in pounds)

$$7.5, 4.5, 10.1, 9.6, and\ 5.5$$

$$Range = 10.1 - 4.5 = 5.6\ pounds$$

Mean Deviation

- Let $x_1, x_2, \ldots, x_n$ be n observations, where $\bar{x}$ is average/ or mean
- $MD = \frac{\sum |x_i - \bar{x}|}{n}$ [for ungrouped data]
- $MD = \frac{\sum f_i \, |x_i - \bar{x}|}{\sum f_i}$ [for grouped data]

Mean Deviation

- For example, $x = 1,2,4,5$.

Mean Deviation

- For example, $x = 1,2,4,5$. Calculate the mean deviation.

$\boldsymbol{x_i}$
1
2
4
5

Mean Deviation

- For example, $x = 1,2,4,5$. Calculate the mean deviation.

$\boldsymbol{x_i}$	$\overline{\boldsymbol{x}}$	$\lvert \boldsymbol{x_i} - \overline{\boldsymbol{x}} \rvert$
1		2
2	$\frac{\sum x_i}{n} = 3$	1
4		1
5		2

$$MD = \frac{\sum |x_i - \bar{x}|}{n} = \frac{6}{4} = 1.5$$

Mean Deviation

Sale (in $)	20-30	30-40	40-50	50-60	60-70
f_i	3	5	12	3	2

Calculate the mean deviation about mean. (Ans: 7.776)

Variance

- Let $x_1, x_2, \dots, x_n$ be n observations

For grouped, $\sigma^2 = \frac{\sum f_i(x_i-\mu)^2}{N}$

- $\sigma^2 = \frac{\sum(x_i-\mu)^2}{N}$ [Population variance (ungrouped)]

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

- $s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$ [Sample variance (ungrouped)]

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$$

For grouped, $s^2 = \frac{\sum f_i(x_i-\bar{x})^2}{n-1}$

Standard Deviation

- Square root of variance

- $\sigma = \sqrt{\frac{\sum(x_i-\mu)^2}{N}}$ [Population standard deviation (ungrouped)]

- $s = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}$[Sample variance (ungrouped)]

Standard Deviation

- For example: weights of 10 newly born babies (in pounds)

$$7.5, 4.5, 10.1, 9.6, 5.5, 6.6, 7.8, 5.9, 6.0, 5.5$$

- a) Calculate the standard deviation

- b) Interpretation

Standard Deviation

x_i	$\bar{x}$	$(x_i - \bar{x})^2$

Standard Deviation

x_i	$\bar{x}$		$(x_i - \bar{x})^2$
7.5	$\frac{\sum x_i}{n} = 6.9$		0.36
4.5			0.57
10.1			10.24
9.6			7.29
5.5			1.96
6.6			0.09
7.8			0.81
5.9			1
6.0			0.81
5.5			1.96

Standard Deviation

- For example: weights of 10 newly born babies (in pounds)
- $$7.5, 4.5, 10.1, 9.6, 5.5, 6.6, 7.8, 5.9, 6.0, 5.5$$
- a) Calculate the standard deviation
-
- b) Interpretation

Standard Deviation

- For example: weights of 10 newly born babies (in pounds)
- $$7.5, 4.5, 10.1, 9.6, 5.5, 6.6, 7.8, 5.9, 6.0, 5.5$$
- a) Calculate the standard deviation
- $s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{30.28}{10-1}} = \sqrt{3.36} = 1.83\ pounds$
- b) Interpretation

The average variation of the weights of the newly born babies from the mean weights is 1.83 pounds.

Standard Deviation (Alternative)

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = 3.36$$

x_i	$\bar{x}$	x_i^2
7.5		56.25
4.5		20.25
10.1		102.01
9.6		92.16
5.5	$\frac{\sum x_i}{n} = 6.9$	30.25
6.6		43.56
7.8		60.84
5.9		34.81
6.0		36
5.5		30.25

Relative measures

- To compare the variation of two or more data sets having different or same units

Relative measures

- To compare the variation of two or more data sets having different or same units

Coefficient of Range
Coefficient of MD
Coefficient of variance
Coefficient of QD

Relative measures

- To compare the variation of two or more data sets having different or same units

Coefficient of Range

Coefficient of MD

Coefficient of variance

Coefficient of QD

Coefficient of variance (CV)

- It is a ratio
- Between standard deviation (SD) and mean
- Expressed as percent

$$\text{General formula of CV} = \frac{SD}{mean} \times 100$$

Coefficient of variance (CV)

- $CV = \frac{\sigma}{\mu} \times 100\,; [For\ population]$

- $CV = \frac{s}{\bar{x}} \times 100\,; [For\ sample]$

Higher CV =
High variation

Lower CV =
Low Variation

More consistent

More uniform

More homogeneous

More stable

Coefficient of variance (CV)

- Section A: Mean = 30 and SD = 4

$$\boxed{CV_A = \frac{4}{30} \times 100 = 13.3\%}$$

- Section B: Mean = 25 and SD = 6

$$CV_B = \frac{6}{25} \times 100 = 24\%$$

- Which section is more consistent in getting final exam mark?

Since, the CV of A is lower than the CV of B.
Thus, section A is more consistent comparatively section B

Thank You