**Measures of Variation (Dispersion)**

It means the measures of the variability or spread among the distributed values of a data set.

**Note: If all values are the same, the dispersion is zero.**

1. **The Range**

The range is the difference between the smallest and largest value in a set of observations, we compute the range as follows:

$$R= x\_{L}-x\_{S}$$

$$Example:2, 5, 7, 10, 70, 10$$

1. **Mean Deviation M.D.**

It is the **average** of different in the **distribution** and the **mean** of that **distribution**. This is the reason why it is commonly called **M.D**.

1. **Un-grouped data (without frequency)**

$$M.D. =\frac{ \sum\_{}^{}\left|x\_{i}-\overbar{x}\right|}{n-1} sample$$

$$M.D. =\frac{ \sum\_{}^{}\left|x\_{i}-μ\right|}{N} population$$

1. **Grouped data**

$$M.D. =\frac{ \sum\_{}^{}f\_{i}\left|x\_{i}-\overbar{x}\right|}{\sum\_{}^{}f\_{i}-1} sample $$

$$M.D. =\frac{ \sum\_{}^{}f\_{i}\left|x\_{i}-μ\right|}{\sum\_{}^{}f\_{i}} population$$

It means the degree of deviation from the mean value; $\left|absolute value\right|$ it does not take the charges $(+, -)$ into the count.

1. **Standard Deviation**

The standard deviation of a set of data is the positive square root of the variance. The standard deviation is $δ$ (**sigma**) in the case of **population**, while it is $S$ in the case of **sample**.

1. **Un-grouped data (without frequency)**

$$S= \sqrt{S^{2}}= \sqrt{\frac{\sum\_{}^{}\left(x\_{i}-\overbar{x}\right)^{2}}{n-1}} sample$$

$$δ= \sqrt{δ^{2}}= \sqrt{\frac{\sum\_{}^{}\left(x\_{i}-μ\right)^{2}}{N}} population$$

$$ $$

1. **Grouped data**

$$S= \sqrt{S^{2}}= \sqrt{\frac{\sum\_{}^{}f\_{i}\left(x\_{i}-\overbar{x}\right)^{2}}{\sum\_{}^{}f\_{i}-1}} sample$$

$$δ= \sqrt{δ^{2}}= \sqrt{\frac{\sum\_{}^{}f\_{i}\left(x\_{i}-\overbar{x}\right)^{2}}{\sum\_{}^{}f\_{i}}} population$$

1. **Variance**

It is a measure of variation used to **measure** the **variance** among the **spread data** from their **mean**. When the values are near from their mean, it means the dispersion is lesser than the case in which they were scattered on a wide range.

1. **Un-grouped data (without frequency)**

$$S^{2}= \frac{\sum\_{}^{}\left(x\_{i}-\overbar{x}\right)^{2}}{n-1} sample $$

$$δ^{2}= \frac{\sum\_{}^{}\left(x\_{i}-μ\right)^{2}}{N} population$$

1. **Grouped data**

$$S^{2}= \frac{\sum\_{}^{}f\_{i}\left(x\_{i}-\overbar{x}\right)^{2}}{\sum\_{}^{}f\_{i}-1} sample $$

$$δ^{2}= \frac{\sum\_{}^{}f\_{i}\left(x\_{i}-μ\right)^{2}}{\sum\_{}^{}f\_{i}} population$$

1. **The Coefficient of Variation**

When **one desires** to compare the **dispersion** in **two sets of data**, however, comparing the two standard deviations may lead to **false** **results**. It may be that the two variables involved are measured in different units. For example, we may wish to know, for a certain population, whether **serum cholesterol levels**, measured in **milligrams per 100 ml**, are more variable than **body weight**, measured in **kilograms**.

$$C.V= \frac{S}{\overbar{x}}×100 sample $$

$$C.V= \frac{δ}{μ}×100 population$$

1. **Standard Error**

The **standard error** is the [standard deviation](http://en.wikipedia.org/wiki/Standard_deviation) of the [sampling distribution](http://en.wikipedia.org/wiki/Sampling_distribution) of a [statistic](http://en.wikipedia.org/wiki/Statistic). The standard error of a statistic depends on the sample size. In general, the larger **sample size** the **smaller the standard error**.

$$S.E= \frac{S.D}{\sqrt{n}}$$

1. **Aberrant value**

Sometimes there is an odd value in our experiment in this case we call it aberrant value. It has a high response or no response and if this value is more than **4 times MD**, it means that this value is an aberrant one, so it must not be taken into our count

**Example**: The age of students at intermediate and preparatory schools were ranged from 12-19 years, but only one of the weights is **40 years**. So, this value should be studied to know whether this is an aberrant value or it must stay and be taken in to the count.

|  |  |
| --- | --- |
| $$Age (g)$$ |  |
| 12 |  |
| 13 |  |
| 14 |  |
| 15 |  |
| 16 |  |
| 17 |  |
| 18 |  |
| 19 |  |
| **40** | **Aberrant value?** |

**Solution**

**1st step**: find **Mean Deviation**

|  |  |  |
| --- | --- | --- |
| $$Weight (x\_{i})$$ | $$x-\overbar{x}$$ | $$\left|x-\overbar{x}\right|$$ |
| 12 | -6.222 | 6.222 |
| 13 | -5.222 | 5.222 |
| 14 | -4.222 | 4.222 |
| 15 | -3.222 | 3.222 |
| 16 | -2.222 | 2.222 |
| 17 | -1.222 | 1.222 |
| 18 | -0.222 | 0.222 |
| 19 | 0.778 | 0.778 |
| 40 | 21.778 | 21.778 |
| $$\overbar{x}=\frac{\sum\_{}^{}x\_{i}}{n}=\frac{164}{9}$$$$\overbar{x}=18.222$$ | $$\sum\_{}^{}\left|x-\overbar{x}\right|=45.11$$ |
| $$M.D= \frac{\sum\_{}^{}\left|x-\overbar{x}\right|}{n-1}= \frac{45.11}{9-1}=\frac{45.11}{8}=5.638$$ |
| $$M.D= 5.638$$$$M.D ×4=5.638×4=22.555$$ |

**2nd step:** if the aberrant value **more than** $(M.D ×4)$ **not** **uses** the aberrant value in the experiment.

**Question 1.** When $\sum\_{}^{}x\_{i} / N=44$, $δ=2.02$, $calculate C.V$.

**Question 2.** A dice was rolled 20 times. On each roll the dice shows a value from 1 to 6. The results have been recorded in the table below, and **find** the **mean**, from a table of data

|  |  |  |
| --- | --- | --- |
| $$Value (x\_{i})$$ | $$Frequency (f\_{i})$$ | $$x\_{i}f\_{i}$$ |
| 1 | 3 | 3 |
| 2 | 5 | 10 |
| 3 | 2 | 6 |
| 4 | 4 | 16 |
| 5 | 3 | 15 |
| 6 | 3 | 18 |
|  | $$\sum\_{}^{}f\_{i}=20$$ | $$\sum\_{}^{}x\_{i}f\_{i}=68$$ |

**Question 3.** Find the value of ($F$) from the following data

|  |  |  |
| --- | --- | --- |
| $$x\_{i}$$ | $$f\_{i}$$ | $$x\_{i}f\_{i}$$ |
| 2 | 1 | 2 |
| 4 | 7 | 28 |
| $$F$$ | 2 | $$2F$$ |
| 8 | 5 | 40 |
|  | $$\sum\_{}^{}f\_{i}=15$$ | $$\sum\_{}^{}x\_{i}f\_{i}=70+2F$$ |

When **mean** of data is = **5.466**

**Question 4.** The tall of students in the class one ranged from **59 – 75 cm** but only one of them was **125 cm**. So, this value should be studied to know whether this is an **aberrant value** or it must stay and be **taken in to the count**.

|  |  |  |
| --- | --- | --- |
| $$Weight (g)$$ | $$x-\overbar{x}$$ | $$\left|x-\overbar{x}\right|$$ |
| **59** | **-16.14** | **260.50** |
| **62** | **-13.14** | **172.66** |
| **65** | **-10.14** | **102.82** |
| **69** | **-6.14** | **37.70** |
| **71** | **-4.14** | **17.14** |
| **75** | **-0.14** | **0.02** |
| **125** | **49.86** | **2486.02** |
| $$\sum\_{}^{}x\_{i}=526$$ | **Aberrant value? Or taken in to the count?** | $$\sum\_{}^{}\left|x-\overbar{x}\right|=3076.86$$ |

### Question 5. Calculating the Mean, Median, and Mode for

**Set A** contains the numbers **2, 2, 3, 5, 5, 7, 8** and

**Set B** contains the numbers **2, 3, 3, 4, 6, 7**.

**Question 6.** The accompanying table shows the scores on a classroom test.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| $$x\_{i}$$ | $$f\_{i}$$ | $$x\_{i}f\_{i}$$ | $$x\_{i}-\overbar{x}$$ | $$\left(x\_{i}-\overbar{x}\right)^{2}$$ | $$f\_{i}\left(x\_{i}-\overbar{x}\right)^{2}$$ |
| **100** | **7** | **700** | **12** | **144** | **1008** |
| **90** | **10** | **900** | **2** | **4** | **40** |
| **80** | **4** | **320** | **-8** | **64** | **256** |
| **70** | **4** | **280** | **-18** | **324** | **1296** |
|  | $$\sum\_{}^{}f\_{i}=25$$ | $$\sum\_{}^{}x\_{i}f\_{i}=2200$$ |  |  | $$\sum\_{}^{}f\_{i}\left(x\_{i}-\overbar{x}\right)^{2}=2600$$ |

What is the **standard deviation** for this set of scores?