**Measures of location (central tendency)**

A **central tendency** (**measure of central tendency**) is a central value or a **typical value** for a probability distribution. It is occasionally called an **average** or just the **center** of the distribution. The most common measures of central tendency are the arithmetic mean, median and the mode.

**Example:** find if you have the following data set

$$x=1, 2, 3, 4, 5, 6, 7$$

$$y=1, 3, 5, 7, 9, 11, 13$$

$$\sum\_{i=1}^{2}x\_{i, } \sum\_{i=2}^{5}x\_{i, } \sum\_{i=5}^{6}x\_{i, } \sum\_{i=1}^{n}x\_{i}^{2}$$

$$(\sum\_{i=1}^{n}x\_{i})^{2} \sum\_{i=1}^{n}x\_{i}y\_{i} \left(\sum\_{i=1}^{n}x\_{i}\right)\left(\sum\_{i=1}^{n}y\_{i}\right)$$

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

$$\sum\_{i=1}^{n}\left(x\_{i}y\_{i}\right)^{2} \sum\_{i=1}^{n}\left(x\_{i}-3\right)+\left(y\_{i}-2\right)$$

$$\sum\_{}^{}\left(x\_{i}-\overbar{x}\right)=0$$

**Mean**

The sum of all measurements divided by the number of observations in the data set, most widely used measure of central location.

1. **Un-grouped data**

In case of frequency is not given, it's simply calculated by taking summation then divide it by total number of observations.

$$\overbar{x}=\frac{\sum\_{}^{}x\_{i}}{n} sample$$

$$μ=\frac{\sum\_{}^{}x\_{i}}{N} population$$

1. **Grouped data**

But when the frequency is given, it could be calculated as follow:

$$\overbar{x}=\frac{\sum\_{}^{}x\_{i}f\_{i}}{\sum\_{}^{}f\_{i}} sample$$

$$μ=\frac{\sum\_{}^{}x\_{i}f\_{i}}{\sum\_{}^{}f\_{i}} population$$

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Midpoint (**$xi$**)** | **Frequency (**$fi$**)** | $$x\_{i}f\_{i}$$ |
| 3.1-3.5 | 3.3 | 5 | 16.5 |
| 3.6-4.0 | 3.8 | 18 | 68.4 |
| 4.1-4.5 | 4.3 | 25 | 107.5 |
| 4.6-5.0 | 4.8 | 27 | 129.6 |
| 5.1-5.5 | 5.3 | 20 | 106 |
| 5.6-6.0 | 5.8 | 5 | 29 |
|  | $$\sum\_{}^{}x\_{i}=27.3$$ | $$\sum\_{}^{}f\_{i}=100$$ | $$\sum\_{}^{}x\_{i}f\_{i}=457$$ |

$$μ=\frac{\sum\_{}^{}x\_{i}f\_{i}}{\sum\_{}^{}f\_{i}}$$

$$μ=\frac{457}{100}=4.57$$

$$\overbar{x}=A+\left[h×\frac{\sum\_{}^{}\left(f\_{i}×d\_{i}\right)}{\sum\_{}^{}f\_{i}}\right]$$

**Where**

$$A=middle value of x\_{i}$$

$$d\_{i}=x\_{i}-A$$

$$h=class interval$$

**Median (middle)**

The value of the middle item in a set of observations which has been arranged in an ascending or descending order of magnitude and is the centermost value in a distribution.

1. **Un-grouped data**
2. **Odd number**

$$Me= \frac{n+1 }{2} sample$$

$$Me= \frac{N+1}{2} population$$

1. **Even number**

$$Me= \frac{n }{2} sample$$

$$Me= \frac{N}{2} population$$

$$Me= \frac{n }{2}+1 sample$$

$$Me= \frac{N}{2}+1 population$$

1. **Grouped data**

$$Me=L+\frac{CI}{f}×\left(\frac{\sum\_{}^{}f\_{i}}{2}-cf\right)$$

**Where**

$Me$ = Median of the distribution

$L$= lower limit of the median class.

$\sum\_{}^{}f\_{i}$ = Summation of total frequencies.

$f$ = frequency of median class.

$cf$ = cumulative frequency before **median class** exactly.

$CI$= class interval.

**Example**: Find the $Me$ of the continuous frequency distribution below

|  |  |  |
| --- | --- | --- |
| **Class** | **Frequency** | **Cumulative Frequency** |
| 5-10 | 80 | 80 |
| 10-15 | 70 | 150 |
| 15-20 | 120 | 270 |
| **20-25** | **210** | 480 |
| 25-30 | 180 | 660 |
| 30-35 | 40 | 700 |
|  | $$\sum\_{}^{}f\_{i}=700$$ | $$\frac{\sum\_{}^{}f\_{i}}{2}=\frac{700}{2}=350$$The greater value from cumulative frequency is **480** (class 20-25) |

**Where**

$Me$ = ?

$L$= 20

$\sum\_{}^{}f\_{i}$ = 700

$f$ = 210

$cf$ = 270

$CI$= 5

$$Me=20+\frac{5}{210}×\left(\frac{700}{2}-270\right)$$

$$Me=21.9$$

**Mode**

The mode of a set of values is that value which occurs most frequently. If all the values are different there is no mode; on the other hand, a set of values may have more than one mode.

$$Mo=L+\left(\frac{f\_{m}-f\_{1}}{\left(2f\_{m}-f\_{1}-f\_{2}\right)}\right)×CI$$

**Where**

$Mo$ = Mode of the distribution

$L$ = lower limit of the modal class.

$f\_{m}$ = frequency of modal class.

$f\_{1}$ = frequency just before the modal class.

$f\_{2}$= frequency just after the modal class.

$CI$ = Class interval.

The approximate value of the mode can be computed using the following relationship:

$$Mode = 3 × Median - 2 × Mean$$

**Example**: Find the $Mo$ of the continuous frequency distribution below

|  |  |  |
| --- | --- | --- |
| **Class** | **Frequency** | The most frequent class number is representing the modal class |
| 5-10 | 80 |
| 10-15 | 70 |
| 15-20 | 120 |
| **20-25** | 210 |
| 25-30 | 180 |
| 30-35 | 40 |
|  | $$\sum\_{}^{}f\_{i}=700$$ |

**Where**

$Mo$ = ?

$L$ = 20

$f\_{m}$ = 210

$f\_{1}$ = 120

$f\_{2}$= 180

$C$ = 5

$$Mo=20+\left(\frac{210-120}{\left(2×210-120-180\right)}\right)×5$$

$$Mo= 23.75$$

**Example**: If you have Plasma Glucose Values (mg/dl) for a Sample of 100 Adults Aged 20–74 Years, find mean, mode and median?

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **70** | 81 | 83 | 87 | 90 | 91 | 94 | 97 | 101 | 106 |
| 75 | 81 | 83 | 87 | 90 | 92 | 95 | 98 | 102 | 107 |
| 77 | 81 | 83 | 88 | 90 | 92 | 95 | 98 | 102 | 109 |
| 78 | 82 | 84 | 88 | 90 | 92 | 95 | 99 | 103 | 111 |
| 78 | 82 | 84 | 89 | 90 | 92 | 95 | 99 | 103 | 112 |
| 78 | 82 | 84 | 89 | 90 | 93 | 95 | 99 | 103 | 114 |
| 80 | 82 | 84 | 89 | 90 | 93 | 96 | 100 | 104 | 115 |
| 80 | 83 | 85 | 89 | 90 | 94 | 96 | 100 | 104 | 116 |
| 80 | 83 | 86 | 89 | 90 | 94 | 97 | 101 | 105 | 117 |
| 80 | 83 | 86 | 90 | 91 | 94 | 97 | 101 | 106 | **117** |

Class number = 1 + 3.322 log n

 = 1 + 3.322 \* log 100

 = 1 + 3.322 \* 2

 = 1 + 6.644

 = 7.644

**Class number = 8**

$$Class Interval =\frac{Range}{No. of Classes}$$

$$Class Interval =\frac{117-70}{7.644}$$

$$Class Interval =\frac{47}{7.644}$$

**Class interval =** 6.14 or 6

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **Frequency**$f\_{i}$ | **Midpoint**$x\_{i}$ | $$x\_{i}f\_{i}$$ |
| 70 – 75 | 2 | 73 | 146 |
| 76 – 81 | 15 | 79 | 1185 |
| 82 – 87 | 17 | 85 | 1445 |
| 88 – 93 | 27 | 91 | 2457 |
| 94 – 99 | 17 | 97 | 1649 |
| 100 – 105 | 13 | 103 | 1339 |
| 106 – 111 | 4 | 109 | 436 |
| 112 – 117 | 5 | 115 | 575 |
|  | $$\sum\_{}^{}f\_{i}=100$$ |  | $$\sum\_{}^{}x\_{i}f\_{i}=9232$$ |

$$μ=\frac{\sum\_{}^{}x\_{i}f\_{i}}{\sum\_{}^{}f\_{i}}$$

$$μ=\frac{9232}{100}$$

$$μ=92.32$$

$$Me=L+\frac{h}{f}×\left(\frac{N}{2}-C\right)$$

**Where**

$Me$ = ?

$L$ = 88

$N$ = 100

$f$ = 27

$C$ = 34

$h$ = 6

$$\frac{\sum\_{}^{}f\_{i}}{2}=\frac{100}{2}=50$$

|  |  |  |
| --- | --- | --- |
| **Class** | **Frequency**$f\_{i}$ | $$Cumulative Frequency$$ |
| 70 – 75 | 2 | 2 |
| 76 – 81 | 15 | 17 |
| 82 – 87 | 17 | 34 |
| 88 – 93 | 27 | 61 |
| 94 – 99 | 17 | 78 |
| 100 – 105 | 13 | 91 |
| 106 – 111 | 4 | 95 |
| 112 – 117 | 5 | 100 |
|  | $$\sum\_{}^{}f\_{i}=100$$ |  |

$$Me=88+\frac{6}{27}×\left(\frac{100}{2}-34\right)$$

$$Me=91.55$$

$$Mo=L+\left(\frac{f\_{m}-f\_{1}}{\left(2f\_{m}-f\_{1}-f\_{2}\right)}\right)×c$$

$Mo$ = ?

$L$ = 88

$f\_{m}$ = 27

$f\_{1}$ = 17

$f\_{2}$=17

$c$ = 6

|  |  |
| --- | --- |
| **Class** | **Frequency** $f\_{i}$ |
| 70 – 75 | 2 |
| 76 – 81 | 15 |
| 82 – 87 | 17 |
| 88 – 93 | 27 |
| 94 – 99 | 17 |
| 100 – 105 | 13 |
| 106 – 111 | 4 |
| 112 – 117 | 5 |
|  | $$\sum\_{}^{}f\_{i}=100$$ |

$$Mo=88+\left(\frac{27-17}{\left(2\*27-17-17\right)}\right)×6$$

$$Mo=91$$

**Question**: Construct a frequency distribution with suitable class interval size of marks obtained by 50 students of a class is given below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 12 | 34 | 43 | 50 | 55 | 62 | 75 |
| 15 | 35 | 44 | 51 | 56 | 63 |  |
| 21 | 36 | 46 | 52 | 56 | 64 |  |
| 23 | 38 | 47 | 52 | 57 | 65 |  |
| 26 | 39 | 47 | 52 | 58 | 65 |  |
| 27 | 41 | 48 | 53 | 59 | 67 |  |
| 30 | 42 | 48 | 54 | 59 | 68 |  |
| 33 | 43 | 50 | 54 | 60 | 72 |  |

Class number = 1 + 3.322 log n

 = 1 + 3.322 \* log 50

 = 1 + 3.322 \* 1.698

 = 1 + 5.643

 = 6.643

**Class number = 7**

$$Class Interval =\frac{Range}{No. of Classes}$$

$$Class Interval =\frac{77-12}{7.644}$$

$$Class Interval =\frac{65}{6.643}$$

**Class interval =** 9.78 or 10

|  |  |
| --- | --- |
| **Class** | **Frequency**$f\_{i}$ |
| 12 – 21 |  |
| 22 – 31 |  |
| 32 – 41 |  |
| 42 – 51 |  |
| 52 – 61 |  |
| 62 – 71 |  |
| 72 – 81 |  |
|  |  |