Axis 4:

Measures of central tendency

Introduction

we focus on the teaching of measures of central tendency, that is, a way of representing a set of data by one number. We often call this the 'average'.

Measures of central tendency help you find the middle, or the average, of a dataset. The 3 most common measures of central tendency are the mode, median, and mean.

Types of Measure of Central Tendency

1- Arithmetic Mean
2- Geometric Mean
3- Harmonic Mean
4- Quadratic Mean
5- Mode
6- Median
7- Quartiles
8- Deciles
9-percentiles

1- Arithmetic Mean:

" A value obtained by dividing the sum of all the observations by the number of observation is called **arithmetic Mean**"

$$Arithmetic\ Mean = \frac{Sum\ of\ \ All\ observation}{Nomber\ of\ \ observation}$$

The following are the formulae of Arithmetic mean:

Methods	Ungrouped data	grouped data
Direct method	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$	$\bar{x} = \frac{\sum_{i=1}^{k} n_i . x_i}{n}$;Here $n = \sum_{i=1}^{k} n_i$
Short cut Method	$\bar{x} = x_0 + \frac{\sum_{i=1}^{n} D_i}{n}$	$\bar{x} = x_0 + \frac{\sum_{i=1}^{k} n_i . D_i}{n}$;Here $n = \sum_{i=1}^{k} n_i$
	Where $D = x_i - x_0$ and x_0 is the provisinal or assumed mean	

2- Geometric Mean:

" The n th root of the product of "n" positive values is called **geometric mean** "

$$Geometric\ Mean = \sqrt[n]{product\ of\ \ "n"\ positive\ values}$$

The following are the formulae of geometric mean:

Methods	Ungrouped data	grouped data
Method one	$M_G = \sqrt[n]{x_1 x_2 ....x_n}$	$M_G = \sqrt[n]{x_1^{n_1} . x_2^{n_2} ....x_k^{n_k}}$;Here $n = \sum_{i=1}^{k} n_i$
Method two	$M_G = Anti\ \log \left(\frac{\sum_{i=1}^{n} \log x_i}{n} \right)$	$M_G = Anti\ \log \left(\frac{\sum_{i=1}^{k} n_i . \log x_i}{n} \right)$;Here $n = \sum_{i=1}^{k} n_i$
$Anti\ \log A = 10^{\log A}$ and $Anti\ \ln A = e^{\ln A}$		

3- Harmonic Mean:

"The reciprocal of the Arithmetic mean of the reciprocal of the values is called ***Harmonic mean****"*

$$Harmonic\ Mean = reciprocal\ of\ \left(\frac{Sum\ of\ \ reciprocal\ of\ \ the\ values}{The\ nomber\ of\ \ values}\right)$$

The following are the formulae of harmonic mean:

	Ungrouped data	grouped data
Harmonic mean	$M_H = \dfrac{n}{\sum_{i=1}^{n}\left(\dfrac{1}{x_i}\right)}$	$M_H = \dfrac{n}{\sum_{i=1}^{k}\left(\dfrac{n_i}{x_i}\right)}$;Here $n = \sum_{i=1}^{k} n_i$

4- Quadratic mean OR root mean square

" The square root of the mean squares (the arithmetic mean of the squares) of the set is called ***Quadratic mean****"*

$$Quadratic\ Mean = square\ root\ of\ \left(\frac{Sum\ of\ \ square\ of\ \ the\ values}{The\ nomber\ of\ \ values}\right)$$

The following are the formulae of Quadrtic mean:

	Ungrouped data	grouped data
Quadratic mean	$M_Q = \sqrt{\dfrac{\sum_{i=1}^{n} x_i^2}{n}}$	$M_Q = \sqrt{\dfrac{\sum_{i=1}^{k} n_i . x_i^2}{n}}$;Here $n = \sum_{i=1}^{k} n_i$

5- Mode:

5-1- Mode in case of Ungrouped Data:

"A value that occurs most frequently in a data is called mode" OR

"if two or more values occur the same number of times but most frequently than the Other values, the there is more than one whole"

"If two or more Values occur the Same number of times but most frequently than the other values, then there is more than one mode"

- *The data having one mode is called uni-modal distribution.*
- *The data having two modes is called bi-modal distribution.*
- *The data having more than two modes is called multi-modal distribution.*

5-2- Mode in case of Discrete Grouped Data:

"A value which has the largest frequency in a set of data is called mode"

5-3- Mode in case of of Continuous Grouped Data:

In case of continuous grouped data, mode would lie in the class that carries the highest frequency.

- *This class is called the modal class.*
- *The formula used to compute the value of mode, is given below:*

$$Mo = x_{Min} + \frac{\Delta_1}{\Delta_1 + \Delta_2} * k_{Mo}$$

Where:

$x_{Min} =$ *Lower class boundary of the Mode class.*

$\Delta_1 =$ *difference between the frequency of the modal class and the frequency of the class preceding the modal class (ignoring signs).*

$\Delta_2 =$ *difference between the frequency of the modal class and the frequency of the class succeeding the modal class (ignoring signs).*

$k_{Mo} =$ *class interval of the distribution.*

6- Median

*"when the observation are arranged in ascending or descending order, then a value, that divides a distribution into equal parts, is called **Median**"*

6-1-Median in case of Ungrouped data

The following are the formulae of median:

Median in case of Ungrouped data	
In this case we first arrange the observations in increasing or decreasing order then we use the following formulae for median	
If "n" is odd	$Median = size\ of \left(\frac{n+1}{2}\right) th\ observation$
If "n" is even	$Median = \frac{size\ of \left\{\left(\frac{n}{2}\right) th + \left(\frac{n}{2}+1\right) th\right\} observation}{2}$

6–2– Median in case of Discrete Grouped Data

in case of Discrete Grouped Data, first we find cumulative frequencies and then use the following formula for Median:

$$Median = size\ of \left(\frac{n}{2}\right) th\ observation;\ here\ n = \sum_{i=1}^{k} n_i$$

6–3– Median in case of Continuous Grouped Data

In continuous grouped data, when we are finding median, we first contrust the class boundaries if the classes are discontinuous. Then we find cumulative frequencies and then we use the following two steps:

- *First we determine the median class using n/2.*
- *When the median class is determined, then the following formula is used to find the value of median. i.e.*

$$Median = x_{Min} + \frac{\frac{n}{2} - N_{-1}^{\uparrow}}{n_{i\,(Median)}} * k_{Median} \qquad ; here\ n = \sum_{i=1}^{k} n_i$$

Where:

x_{Min} = *Lower class boundary of the median class.*

k_{Median} = *Width of the median class.*

$n_{i\,(Median)}$ = *Frequency of the median class.*

$N_{-1}^{\uparrow}$ = *Cumulative Frequency of the class preceding the median class.*

7- Quartiles:

"when the observation are arranged in increasing order then the values, that divide the whole data into four (4) equal parts, are called ***Quartiles****"*

These values are denoted by Q_1, Q_2 *and* Q_3. *It is to be noted that 25% of the data falls below* Q_1, *50% of the data falls below* Q_2 *and 75% of the data falls below* Q_3.

– Quartiles in case of Continuous Grouped Data

In continuous grouped data, when we are finding Quartiles, we first contrust the class boundaries if the classes are discontinuous. Then we find cumulative frequencies and then we use the following two steps:

- *First we determine the Quartile class using:* $j.\left(\frac{n}{4}\right);\quad here\ \ j = 1.2.3$.
- *When the Quartile class is determined, then the following formula is used to find the value of Quartile. i.e.*

$$Q_j = x_{Min} + \frac{j.\left(\frac{n}{4}\right) - N_{-1}^{\uparrow}}{n_{i\,(Q_j)}} * k_{Q_j} \qquad ; here\ n = \sum_{i=1}^{k} n_i\ \ and\ \ j = 1, 2, 3$$

Where:

x_{Min} =Lower class boundary of the Quartile class.

k_{Q_j} = Width of the Quartile class.

$n_{i(Q_j)}$ =Frequency of the Quartile class.

$N_{-1}^{\uparrow}$ = Cumulative Frequency of the class preceding the Quartile class.

8- Deciles:

*"when the observation are arranged in increasing order then the values, that divide the whole data into ten (10) equal parts, are called **Deciles**".*

These values are denoted by $D_1, D_2, ..., D_9$. *It is to be noted that 10% of the data falls below* D_1, *20% of the data falls below* D_2, *..., and 90% of the data falls below* D_9.

– Deciles in case of Continuous Grouped Data

In continuous grouped data, when we are finding Deciles, we first contrust the class boundaries if the classes are discontinuous. Then we find cumulative frequencies and then we use the following two steps:

- *First we determine the Decile class using:* $j.\left(\frac{n}{10}\right)$; *here* $j = 1, 2......, 9$.
- *When the Quartile class is determined, then the following formula is used to find the value of Decile. i.e.*

$$D_j = x_{Min} + \frac{j.\left(\frac{n}{10}\right) - N_{-1}^{\uparrow}}{n_{i(D_j)}} * k_{D_j} \qquad ; here\ n = \sum_{i=1}^{k} n_i \ \ and \ \ j = 1, 2...., 9$$

Where:

x_{Min} =Lower class boundary of the Decile class.

k_{D_j} = Width of the Decile class.

$n_{i(D_j)}$ =Frequency of the Decile class.

$N_{-1}^{\uparrow}$ = Cumulative Frequency of the class preceding the Decile class.

9- Percentiles:

"when the observation are arranged in increasing order then the values, that divide the whole data into hundred (100) equal parts, are called ***percentiles****".*

These values are denoted by $P_1, P_2, \ldots, P_{99}$. *It is to be noted that 1% of the data falls below* P_1, *2% of the data falls below* $P_2, \ldots,$ *and 99% of the data falls below* P_{99}.

– Percentiles in case of Continuous Grouped Data

In continuous grouped data, when we are finding Percentiles, we first contrust the class boundaries if the classes are discontinuous. Then we find cumulative frequencies and then we use the following two steps:

- *First we determine the Percentile class using:* $j.\left(\frac{n}{100}\right)$; $here\ \ j = 1, 2 \ldots.., 99$.
- *When the Quartile class is determined, then the following formula is used to find the value of Decile. i.e.*

$$P_j = x_{Min} + \frac{j.\left(\frac{n}{100}\right) - N_{-1}^{\uparrow}}{n_{i(P_j)}} * k_{P_j} \qquad ; here\ n = \sum_{i=1}^{k} n_i \ \ and \ \ j = 1, 2 \ldots.., 99$$

Where:

x_{Min} = *Lower class boundary of the Percentile class.*

k_{P_j} = *Width of the Percentile class.*

$n_{i(P_j)}$ = *Frequency of the Percentile class.*

$N_{-1}^{\uparrow}$ = *Cumulative Frequency of the class preceding the Percentile class.*

- $Median = Q_2 = D_5 = P_{50}$
- $Q_1 = P_{25} \quad and \quad Q_3 = P_{75}$
- $D_1 = P_{10} \quad and \quad D_2 = P_{20}.....................and\ D_9 = P_{90}$

Uses of Averages in Different Situations

- $\bar{x}$ *is an appropriate average for all the situations where there are no extreme values in the data.*
- M_G *is an appropriate average for calculating average percent increase in sales, population, production, etc. It is one of the best averages for the construction of index numbers.*
- M_H *is an appropriate average for calculating the average rate of increase of profits of a firm or finding average speed of a journey or the average price at which articles are sold.*
- *Mode: is an appropriate average in case of qualitative data e.g. the opinion of an average person; he is probably referring to the most frequently expressed opinion which is the modal opinion.*
- *Median is an appropriate average in a highly skewed distribution e.g. in the distribution of wages, incomes etc.*