UNIT 3 INTRODUCTION TO MEASURES OF CENTRAL TENDENCY*

Structure

* Prof. Suhas Shetgovekar, Faculty, Discipline of Psychology, School of Social Sciences, IGNOU, New Delhi

3.0 OBJECTIVES

After reading this unit, you will be able to:

- explain the concept of central tendency of data;
- describe the different measures of central tendency;
- discuss the properties, advantages and limitations of mean, median and mode; and
- compute measures of central tendency for ungrouped and grouped data.

3.1 INTRODUCTION

Suppose you have data, for instance, marks in psychology obtained by students in 12th standard and you want to analyse it statistically, what statistical techniques will you employ? You can of course organise the data with the help of classification and tabulation that we discussed in the previous Unit and the data can also be graphically represented. But if you want to further analyse the data then you can compute the average marks obtained by the whole class or find the midpoint for marks above and below which will lie half of the students or you can also find out most frequent marks obtained by the students. The techniques you are employing here are mean, median and mode. These are called measures of central tendency and can be categorised under descriptive statistics.

In the previous unit, we discussed about classification, tabulation and also graphical representations of data. In the present unit, we will discuss the measures of central tendency, viz., mean, median and mode. We will not only understand what these techniques are, but will also focus on their properties, advantages and limitations. Further, we will also learn how to compute mean, median and mode for grouped and ungrouped data.

3.2 CONCEPT OF CENTRAL TENDENCY OF DATA

Measures of central tendency provides a single value that indicates the general magnitude of the data and this single value provides information about the characteristics of the data by identifying the value at or near the central location of the data (Bordens and Abbott, 2011). King and Minium (2013) described measures of central tendency as a summary figure that helps in describing a central location for a certain group of scores. Tate (1955, page 78) defined measures of central tendency as “a sort of average or typical value of the items in the series and its function is to summarise the series in terms of this average value”.

The main functions of measures of central tendency are as follows:

1) They provide a summary figure with the help of which the central location of the whole data can be explained. When we compute an average of a certain group we get an idea about the whole data.

2) Large amount of data can be easily reduced to a single figure. Mean, median and mode can be computed for a large data and a single figure can be derived.

3) When mean is computed for a certain sample, it will help gauge the population mean.

4) The results obtained from computing measures of central tendency will help in making certain decisions. This holds true not only to decisions with regard to research but could have applications in varied areas like policy making, marketing and sales and so on.

5) Comparison can be carried out based on single figures computed with the help of measures of central tendency. For example, with regard to performance of students in mathematics test, the mean marks obtained by girls and the mean marks obtained by boys can be compared.

A good measure of central tendency needs to have the following characteristics:

1) The definition of the central tendency needs to be adequately specified and should be clear. It should not be subject to varied interpretations and needs to be unaffected by any individual bias. The definition should be rigid so that a stable value is obtained that represents the data.

2) The measure of central tendency should be easy to understand and easy to compute. It should not involve elaborate mathematical calculations.

3) For the value obtained from the computation of measures of central tendency to be representative of the data, the whole data needs to be computed.

4) The data needs to be collected from a sample that truly represents the population. The sample thus needs to be randomly selected.

5) The measure of central tendency needs to display sampling stability and should not be affected by any fluctuations in the sample. For example, if two different researchers obtain a representative sample from a same population, the means computed by them for their respective sample should display least variation.

6) The measure of central tendency should not be affected by outliers. Outliers are extreme values in data or distribution.

7) The measure of central tendency should render itself to further mathematical computations.

Check Your Progress I

1) Define measures of central tendency.

....

....

....

....

....

....

2) List the functions of measures of central tendency.

....

....

....

....

....

....

3.3 DIFFERENT MEASURES OF CENTRAL TENDENCY

As the concept of central tendency is now clear, we will now proceed to discuss the three measures of central tendency. The three measures of central tendency that we will be discussing are:

1) Mean or Arithmetic mean
2) Median
3) Mode

In this section of the Unit, we will try to understand these concepts and then in the next section we will be focusing on the properties, advantages and limitations of each of these measures.

3.3.1 Mean or Arithmetic Mean

Mean for sample is denoted by symbol 'M or $\overline{x}$ ('x-bar')' and mean for population is denoted by 'μ' (mu). It is one of the most commonly used measures of central tendency and is often referred to as average. It can also be termed as one of the most sensitive measure of central tendency as all the scores in a data are taken in to consideration when it is computed (Bordens and Abbott, 2011). Further statistical techniques can be computed based on mean, thus, making it even more useful.

Mean is a total of all the scores in data divided by the total number of scores. For example, if there are 100 students in a class and we want to find mean or average marks obtained by them in a psychology test, we will add all their marks and divide by 100, (that is the number of students) to obtain mean.

3.3.2 Median

Median is a point in any distribution below and above which lie half of the scores. Median is also referred to as P_{50} (King and Minium, 2008). The symbol for median is 'M_d'. As stated by Bordens and Abbott (2011, page 411), 'median is the middle score in an ordered distribution'. If we take the example discussed earlier of the marks obtained by 100 students in a psychology test, these marks are to be arranged in an order, either ascending or descending. The middle score in this distribution is then identified as median. Though this would seem easy for an odd number of scores, in case of even number of scores a certain procedure is followed that will be discussed when we learn how to compute median later in this unit.

3.3.3 Mode

Mode is denoted by symbol 'M_o'. Mode is the score in a distribution that occurs most frequently. Taking the example of the marks obtained by a group of 100 students in psychology test discussed earlier, if out of these 100 students, 10 students obtained 35 marks. 35 is thus, most frequently occurring value and will be termed as mode. Certain distributions can be bimodal as well, where there are two modes. For instance if there were other 10 students in this group of 100 students, who secured 47 marks, 47 is the value that is occurring as frequently as 35 and thus, will be termed as mode along with 35. In a similar way, when there are three modes, the term used is trimodal and when there are four or more modes, we use the term multimodal.

Though if the scores in a distribution greatly vary then it is possible that there is no mode. Mode as such does not provide an adequate characterisation of the distribution because it just takes in to consideration the most frequent scores and other scores are not considered.

How to choose a measure of central tendency?

The choice of a measure of central tendency will depend on first of all, the scales of measurement that we discussed in the first unit. For nominal scales one can compute mode but not mean or median. For example, in case of males and females, the males can be coded as 1 and females can be coded as 2 (or vice versa) in such a case, we can compute frequently occurring score, that will provide us information whether there are more males or more females. However, it is not possible to compute mean or median. With regard to ordinal scale, median or mode can be used. For example, if we rank the students basedon their performance in mathematics test, it is possible to find median belowand above which lie half of the ranks. Mode can also be computed if morethan one student gets same rank. With regard to interval scale and ratio scalemean can be computed.

Yet another aspect that is important while making a choice with regard to which measure of central tendency to use is, whether the data is normally distributed or not. If the data is normally distributed we will compute mean and if it is not normally distributed, we will compute median or mode. This is because mean may not adequately represent the data when the data is not normally distributed. We will discuss normal distribution in detail in the last unit (unit 8) of this course.

Check Your Progress II

1) Describe mean, median and mode.

Measure	Description	Example
Mean		

Median		
Mode		

2) How to choose a measure of central tendency?

....

....

....

....

....

3.4 PROPERTIES, ADVANTAGES AND LIMITATIONS OF MEAN, MEDIAN AND MODE

Let us now discuss the properties, advantages and limitations of mean, median and mode.

3.4.1 Properties of Mean

1) Mean is sensitive to the actual position of each and every score in a distribution and if another score is included in the distribution, then the mean or average of that distribution will change. For example, mean of the scores 5, 4 ,6 ,3, 2 is 4 [We got the value 4 by adding 5+4+6+3+2= 20 and then dividing it by 5, that is the total number of scores (N)]. But if we change the scores to 5, 4, 6, 3, 2, 8, the mean will be 4.67 [We got the value 4.67 by adding 5+4+6+3+2+8= 28 and then dividing it by 6, that is the total number of scores (N)]

2) Mean denotes a balance point of any distribution and the total of positive deviations from the mean is equal to the negative deviations from the mean (King and Minium, 2008).

3) Mean is especially effective when we want the measure of central tendency to reflect the sum of the scores.

3.4.2 Advantages of Mean

1) The definition of mean is rigid which is a quality of a good measure of central tendency.

2) It is not only easy to understand but also easy to calculate.

3) All the scores in the distribution are considered when mean is computed.

4) Further mathematical calculations can be carried out on the basis of mean.

5) Fluctuations in sampling are least likely to affect mean.

3.4.3 Limitations of Mean

1) Outliers or extreme values can have an impact on mean.

2) When there are open ended classes, such as 10 and above or below 5, mean cannot be computed. In such cases median and mode can be computed. This is mainly because in such distributions mid point cannot be determined to carry out calculations.

3) If a score in the data is missing or lost or not clear, then mean cannot be computed unless mean is computed for rest of the data by not considering the lost score and dropping it all together.

4) It is not possible to determine mean through inspection. Further, it cannot be determined based on a graph.

5) It is not suitable for data that is skewed or is very asymmetrical as then in such cases mean will not adequately represent the data.

3.4.4 Properties of Median

1) When compared to mean, median is less sensitive to extreme scores or outliers.

2) When a distribution is skewed or is asymmetrical median can be adequately used.

3) When a distribution is open ended, that is, actual score at one end of the distribution is not known, then median can be computed.

3.4.5 Advantages of Median

1) The definition of median is rigid which is a quality of a good measure of central tendency.

2) It is easy to understand and calculate.

3) It is not affected by outliers or extreme scores in data.

4) Unless the median falls in an open ended class, it can be computed for grouped data with open ended classes.

5) In certain cases it is possible to identify median through inspection as well as graphically.

3.4.6 Limitations of Median

1) Some statistical procedures using median are quite complex. Computation of median can be time consuming when large data is involved because the data needs to be arranged in an order before median is computed.

2) Median cannot be computed exactly when an ungrouped data is even. In such cases, median is estimated as mean of the scores in the middle of the distribution.

3) It is not based on each and every score in the distribution.

4) It can be affected by sampling fluctuations and thus can be termed as less stable than mean.

3.4.7 Properties of Mode

1) Mode can be used with variables that can be measured on nominal scale.

2) Mode is easier to compute than mean and media. But it is not used often because of lack of stability from one sample to another and also because a single set of data may possibly have more than one mode. Also, when there is more than one mode, then the modes cannot be termed to adequately measure central location.

3) Mode is not affected by outliers or extreme scores.

3.4.8 Advantages of Mode

1) It is not only easy to comprehend and calculate but it can also be determined by mere inspection.

2) It can be used with quantitative as well as qualitative data.

3) It is not affected by outliers or extreme scores.

4) Even if a distribution has one or more than one open ended classe(s), mode can easily be computed.

3.4. 9 Limitations of Mode

1) It is sometimes possible that the scores in the data vary from each other and in such cases the data may have no mode.

2) Mode cannot be rigidly defined.

3) In case of bimodal, trimodal or multimodal distribution, interpretation and comparison becomes difficult.

4) Mode is not based on the whole distribution.

5) It may not be possible to compute further mathematical procedures based on mode.

6) Sampling fluctuations can have an impact on mode.

Check Your Progress II

1) List the properties of mean.

....

....

....

....

....

2) List the advantages of median.

....

....

....

....

....

3) List the limitations of mode.

....

....

....

....

....

3.5 COMPUTATION OF MEASURES OF CENTRAL TENDENCY IN UNGROUPED AND GROUPED DATA

Now as we have developed a fair idea about the three measures of central tendency, we will move on to learn how to compute them. While computing each of these measures, we will do so for ungrouped and grouped data. Ungrouped and grouped data are explained as follows:

Ungrouped data: Any data that has not been categorised in any way is termed as an ungrouped data. For example, we have an individual who is 25 years old, another who is 30 years old and yet another individual who is 50 years old. These are independent figures and not organised in any way, thus they are ungrouped data.

Grouped data: A data that is categories or organised is termed as grouped data. Mainly such data is organised in frequency distribution. For example, we can have age range 26- 30 years, 31- 35 years, 36- 40 years and so on. Grouped data are convenient especially when the data is large.

3.5.1 Computation of Mean for Ungrouped Data

The formula for computing mean for ungrouped data is

$M = \Sigma X/ N$

Where,

M = Mean

ΣX= Summation of scores in the distribution

N = Total number of scores.

Let us now compute mean with the help of an example

The scores obtained by 10 students on psychology test are as follows:

58 34 32 47 74 67 35 34 30 39

Step 1: In order to obtain mean for the above data we will first add the marks to obtain ΣX:

58+ 34+ 32+ 47+ 74+ 67+ 35+ 34+ 30+ 39 = 450

Step 2: Now using the formula, we will compute mean

$M = \Sigma X/ N$

ΣX= 450 , N= 10 (Total number of students)

Thus,

M= 450/ 10 = 45

Thus, the mean obtained for the above data is 45

3.5.2 Computation of Mean for Grouped Data

The formula for computing mean for grouped data is

$M = \Sigma fX/ N$

Where,

M= Mean

$\sum$= Summation

X= Midpoint of the distribution

f = The respective frequency

N = Total number of scores.

Let us now compute mean with the help of an example.

A class of 30 students were given a psychology test and the marks obtained by them were categorised in to six categories. The lowest marks obtained were 10 and highest marks obtained were 35. A class interval of 5 was employed. The data is given as follows:

Marks	Frequencies (*f*)	Midpoint (X)	*f*X
35- 39	5	37	185
30-34	7	32	224
25-29	5	27	135
20-24	6	22	132
15-19	4	17	68
10- 14	3	12	36
	N= 30		**ΣfX = 780**

The steps followed for computation of mean with grouped data are as follows:

Step 1: The data is arranged in a tabular form with marks grouped in categories with class interval of 5.

Step 2: Once the categories are created, the marks are entered under frequency column based on which category they fall under.

Step 3: The midpoints of the categories are computed and entered under X.

Step 4: *f*X is obtained by multiplying the frequencies and midpoints for each category.

Step 5: *f*X for all the categories are added to obtain ΣfX, in case of our example it is obtained as 780

Step 6: The formula M = Σ*f*X/ N is used, N is equal to 30.

M = Σ*f*X/ N

M = 780/ 30= 26

Thus, the mean obtained is 26

3.5.3 Computation of Mean by Shortcut Method (with Assumed mean)

In certain cases data is very large and it is not possible to compute each *f*X. In such situations, a short cut method with the help of assumed mean can be computed. A real mean can thus be computed with application of correction.

The formula is

M =AM+(Σfx′/N×i)

Where,

AM= Assumed mean,

Σ = Summation

i = Class interval

$x' = \{(X - AM)/ i\}$, X the midpoint of the scores in the interval

f = the respective frequency of the midpoint

N = The total number of frequencies or students.

Let us discuss the steps followed for computation of mean with the help of an example given below:

Class Intervals (Marks)	Frequencies (f)	Midpoint (X)	$x' = \{(X - AM)/ i\}$	fx'
35- 39	5	37	3	15
30-34	7	32	2	14
25-29	5	27	1	5
20-24	6	22	0	0
15-19	4	17	-1	-4
10- 14	3	12	-2	-6
	N= 30			**$\Sigma fx' = 24$**

Step 1: We will assume mean (AM) as 22.

Step 2: Difference is obtained between each of the midpoints and the assumed mean and then the same is divided by 'i' that is the class interval (5 in this case), these are then entered under column with heading $x' = \{(X - AM)/ i\}$. The x' for 22 will be 0.

Step 3: Frequency (f) is then multiplied with x' to obtain fx'.

Step 4: All fx' are added to obtain $\Sigma fx'$, in the present example it is 24.

Step 5: The formula for mean is now applied

$M = AM + (\Sigma fx'/N \times i)$

$M = 22 + (24/30 \times 5)$

$= 22 + 4 = 26$

Thus, mean is obtained as 26.

And if you refer to the mean obtained by the direct method and mean obtained with the shortcut method, the mean is the same, that is 26.

3.5.4 Computation of Median for Ungrouped Data

With regard to computation of median for ungrouped data, different procedures are followed for data that is odd and data that is even.

3.5.4.1 Odd Data: When the data is odd the median is computed in the following manner:

Data: 58 34 32 47 74 67 35 34 30 (N= 9)

Step 1: First the data is to be arranged in either ascending or descending order. We will arrange the data in ascending order and it will look like this:

30 32 34 34 35 47 58 67 74

Step 2: The following formula is then used to compute Median:

M_d = (N +1)/ 2th score

Thus (9 +1)/2= 10/2 =5th item

In our data the 5th item is 35, that is the median of this data.

3.5.4.2 Even data: When the data is even, the median is computed in the following manner:

58 34 32 47 74 67 35 34 30 39 (N= 10)

Step 1: First the data is to be arranged in either ascending or descending order. We will arrange the data in ascending order and it will look like this:

30 32 34 34 35 39 47 58 67 74

Step 2: The following formula is used to compute median:

M_d = (N/2)th score + [(N/2)th score + 1]/ 2

The (N/2)th score is the 5th score, that is 35.

The (N/2)th score + 1 is the 6th score, that is 39.Thus 35 +39/ 2= 37

The median thus obtained is 37.

3.5.5 Computation of Median for Grouped Data

The formula used for computation of median for grouped data is as follows:

$$M_d = L + \left[(N/2) - F / f_m\right] \times i$$

Where,

L = The lower limit of the median class

N = Total of all the frequencies

F = Sum of frequencies before the median class

f_m = frequency within the interval upon which the median falls

i = class interval.

Let us discuss the steps followed for computation of median with the help of the example given below:

Class Intervals (Marks)	Frequencies (*f*)
35- 39	5
30-34	7
25-29	**5**
20-24	6
15-19	4
10- 14	3
	N= 30

The steps in computing median for grouped data are as follows:

Step 1: The first step is to compute N/2, that is 30/2 so that we obtainone half of the scores in the data (15 in this case).

Step 2: As the scores are even in number (N= 30), the median should fall between 15th and 16th score. Whether we add the frequencies from above (5+7+5= 17) or from below (3+4+6+5= 18), the median will fall in the class interval 25-29. Further L that is the lower limit of the median class can also be mentioned. As the median class is 25-29, its lower limit will be 24.5.

Step 3: Compute F, that is sum of frequencies before the median class. In our example it would be 3 +4 +6 = 13

Step 4: f_m is computed. It is the frequency within the interval upon which the median falls. In the present example the median class interval is 25-29 and the frequency for this class interval is 5. So f_m is 5.

Step 5: The values can now be put in the formula to obtain the median

$$\mathbf{M_d = L + [(N/2) - F/f_m] \times i}$$

$= 24.5 + [(30/2) - 13/5] \times 5$

$= 24.5 + [15 - 13/5] \times 5$

$= 24.5 + [2/5] \times 5$

$= 24.5 + 10/5$

$= 24.5 + 2$

$= 26.5$

Thus, the median obtained is 26. 5. And it falls in the median class interval 25-29.

3.5.6 Computation of Mode for Ungrouped Data

Let us now learn how to compute mode for an ungrouped data with the help of the following example:

58 34 32 47 74 67 35 34 30 39

The mode can be calculated in simple manner by just counting the scores that appears maximum number of times in the data. In our example, the score occurring maximum number of times is 34, that occurs twice. Thus the mode is 34.

3.5.7 Computation of Mode for Grouped Data

There are two methods by which mode for grouped data can be computed:

3.5.7.1 First Method

The first method is by using the following formula

Mode=3Mdn−2M

Where,

Mdn = Median

M= Mean

Let us now compute mode with the help of the following example:

Class Intervals (Marks)	Frequencies (*f*)	Midpoint (X)	*f*X
50- 59	5	54.5	272.5
40- 49	7	44.5	311.5
30- 39	8	34.5	276
20- 29	10	24.5	245
10- 19	15	14.5	217.5
0- 9	5	4.5	54.522.5
	N= 50		**ΣfX = 1345**

The formula M = Σ*f*X/ N is used, N is equal to 50.

Step 1: Compute mean

M = Σ*f*X/ N

M = 1370/ 50= 26.9

Step 2: Compute median

$$M_d = L + [(N/2) - F/f_m] \times i$$

$= 19.5 + [(50/2) - 20/10] \times 10$

$= 19.5 + [25 - 20/10] \times 10$

$= 19.5 + [5/10] \times 10$

$= 19.5 + 5$

$= 24.5$

Step 3: Let us now use these values in our formula and compute mode

$\mathbf{M_o = 3Mdn - 2M}$

$M_o = 3 \times 24.5 - 2 \times 26.9$

$= 73.5 - 53.8$

$= 19.7$

Thus the mode computed is 19.7

Also we can make one observation here that the mean obtained for our example is 26.9 the median is 24.5 and the mode is 19.7. All the three values are not close to each other indicating that the distribution of the data may not be normal as the values do not fall in the central area of the distribution. If the values of mean, median and mode were similar, then we could have said that the data is normally distributed.

3.5.7.2 Second Method

In the second method of computing mode for grouped data the following formula is used:

$$M_o = L + [d_1 / d_1 + d_2] \times i$$

Where,

L= Lower limit of the class interval in which the mode may lie, called as modal class

i = Class interval

d_1 = difference between frequencies of modal class and class interval below it.

d_2 = difference between frequencies of modal class and class interval above it,

Let us discuss the steps followed for computation of mode with the help of the example given below:

Class Intervals (Marks)	Frequencies (*f*)
35- 39	5
30-34	**7**
25-29	5
20-24	6
15-19	4
10- 14	3
	N= 30

Step 1: The mode is most likely to fall in the the class intervals 30-34 as that has the highest frequencies (7). Thus this is our modal class and the lower limit of the same (L) will be 29.5.

Step 2: The class interval (i) for this example is 5.

Step 3: Compute d_1, that is, difference between frequencies of modal class and class interval below it and d_2, that is, difference between frequencies of modal class and class interval below it.

$d_1 = f_m - f_{m-1}$

$d_2 = f_m - f_{m+1}$

Where,

f_m = the frequency of the modal class (7 in case of our example).

f_{m-1} = the frequency of the class interval below the modal class (5 in case of our example).

f_{m+1} = the frequency of the class interval above the modal class (5 in case of our example).

Thus, d_1= 7-5 = 2 and d_2= 7-5 = 2 in case of our example.

Step 4: Now let us compute mode with the help of the formula

$$\mathbf{M_o = L + [d_1/\ d_1 + d_2]\ x\ i}$$

$$M_o = 29.5 + [2/2+2] x\ 5$$

$$= 29.5 + 2/4 \times 5$$

$$= 29.5 + 10/4$$

$$= 29.5 + 2.5$$

$$= 32$$

Thus, the mode obtained is 32.

Check Your Progress IV

1) Compute mean, median and mode for the following data:

23, 34, 43, 65, 67, 67, 78, 65, 43, 34, 45, 33, 23, 67, 60 (N= 15)

2) Compute mean for the following data:

Class Intervals (Marks)	**Frequencies (*f*)**
50- 59	4
40- 49	5
30- 39	6
20- 29	5
10- 19	5
1- 9	5
	N= 30

3.6 LET US SUM UP

In the present unit, we discussed the concept of central tendency. The measures of central tendency was explained as summary figures that help in describing a central location for a certain group of scores. It was further explained as providing information about the characteristics of the data by identifying the value at or near the central location of the data. The functions of measures of tendency besides the characteristics of good measures of central tendency were also discussed. Further, the unit focused on the three measures of central tendency, namely, mean, median and mode. Mean is a total of all the scores in data divided by the total number of scores. It is one of the most frequently used measure of central tendency and is often referred to as an average. It can also be termed as one of the most sensitive measure of central tendency as all the scores in a data are taken in to consideration when it is computed. Median is the middle score in an ordered distribution. Median is a point in any distribution below and above which lie half of the scores. Mode is the score in a distribution that occurs most frequently. Certain distributions are bimodal, where there are two modes.When there are three modes, the term used is trimodal and when there are four or more modes, we use the term multimodal. Though, if the scores in a distribution greatly vary, then it is possible that there is no mode. The properties, advantages and limitations of mean, median and mode were also discussed in detail. Further, the computation of each of these measures of central tendency was also discussed for both ungrouped and grouped data with stepwise explanation.

3.7 REFERENCES

Bordens, K. S. and Abbott, B. B. (2011). *Research Design and Methods: A Process Approach*. New Dekhi:McGraw Hill Education(India) Private Limited.

King, Bruce. M; Minium, Edward. W. (2008). *Statistical Reasoning in the Behavioural Sciences*. Delhi: John Wiley and Sons, Ltd.

Mangal, S. K. (2002). *Statistics in Psychology and Education*. new Delhi: Phi Learning Private Limited.

Minium, E. W., King, B. M., & Bear, G. (2001). *Statistical Reasoning in Psychology and Education*. Singapore: John-Wiley.

Mohanty, B and Misra, S. (2016). *Statistics for Behavioural and Social Sciences*. Delhi: Sage.

Tate, M. W.(1955). *Statistics in Education*. New York: Macmillion Co.

Veeraraghavan, V and Shetgovekar, S. (2016). *Textbook of Parametric and Nonparametric Statistics*. Delhi: Sage.

3.8 KEY WORDS

Measures of Central Tendency: Measures of central tendency can be explained as a summary figure that helps in describing a central location for a certain group of scores.

Mean: Mean is a total of all the scores in data divided by the total number of scores.

Median: Median is a point in any distribution below and above which lie half of the scores.

Mode: Mode is the score in a distribution that occurs most frequently.

3.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) Define Measures of Central Tendency

Measures of central tendency can be defined as a summary figure that helps in describing a central location for a certain group of scores. It is a value that determines the general magnitude of a distribution.

1) List the functions of measures of central tendency.

   a) They provide a summary figure with the help of which the central location of the whole data can be explained.

   b) The large amount of data can be easily reduced to a single figure.

   c) When mean is computed for a certain sample, it will help us gain idea about the population mean.

   d) The results obtained from computing measures of central tendency will help a researcher make certain decisions.

   e) Comparison can be carried out with the help of the single figures computed with the help of measures of central tendency.

Check Your Progress II

1) Describe mean, median and mode with suitable examples.

Measure	Description	Example
Mean	Mean is a total of all the scores in data divided by the total number of scores. It is one of the most often used measures of central tendency and is often referred to as average. It can also be termed as one of the most sensitive measures of central tendency as all the scores in a data are taken in to consideration when it is computed.	Scores on Job Satisfaction obtained by 5 employees 23, 34, 54, 34, 22 (N= 5) Thus Mean would be 23 +34 +54 + 34 +22= 167 Thus 167/5= 33.4
Median	Median is the middle score in an ordered distribution. Median is a point in any distribution below and above which lie half of the scores.	In above example, the data is arranged in ascending order 22, 23, 34, 34, 54 Median thus is 34
Mode	Mode is the score in a distribution that occurs most frequently. Certain distributions can be bimodal as well, where there are two modes.When there are three modes, the term used is trimodal and when there are four or more modes, we use the term multimodal. Though if the scores in a distributions greatly vary then it is possible that there is no mode.	In above example, 23, 34, 54, 34, 22 Mode is 34 that occurs twice

2) How to choose which measure of central tendency to use?

Choice of measure of central tendency will depend on the scales of measurement and also whether the data is normally distributed or not.

Check Your Progress III

1) List the properties of mean

   a) Mean is sensitive to the actual position of each and every score in a distribution and if another score is included in the distribution, then the mean or average of that distribution will change.

   b) Mean denotes a balance point of any distribution and the total of positive deviations from the mean is equal to the negative deviations from the mean.

   c) Mean is especially effective when we want the measure of central tendency to needs to reflect the sum of the scores.

2) List the advantages of median.

   a) The definition of median is rigid which is a quality of a good measure of central tendency.

   b) It is easy to understand and calculate.

   c) It is not affected by outliers or extreme scores in data.

   d) Unless the median falls in an open ended class, it can be computed for grouped data with open ended classes.

   e)I n certain cases it is possible to identify median through inspection as well as graphically.

3) List the limitations of mode.

   a) It is sometimes possible that the scores in the data vary from each other and in such cases the data may have no mode.

   b) Mode cannot be rigidly defined.

   c) In case of bimodal, trimodal or multimodal distribution, interpretation and comparison becomes difficult.

   d) Mode is not based on the whole distribution.

   e) It may not be possible to compute further mathematical procedures based on mode.

   f) Sampling fluctuations can have an impact on mode.

Check Your Progress IV

1) Compute mean, median and mode for the following data:

   23, 34, 43, 65, 67, 67, 78, 65, 43, 34, 45, 33, 23, 67, 60 (N= 15)

   Mean = 49.8, Median = 45, Mode: 67

2) Compute mean for the following data:

Class Intervals (Marks)	Frequencies (*f*)
50- 59	4
40- 49	5
30- 39	6
20- 29	5
10- 19	5
1- 9	5
	N= 30

Mean = 28.83

3.10 UNIT END QUESTIONS

1) Discuss the concept of measures of central tendency with a focus on characteristics of a good measure of central tendency.

2) Explain the properties of mean, median and mode.

3) Discuss the limitations of mean, median and mode.

4) Compute mean, median and mode for the following data:

   44, 32, 34, 34, 45,54, 56, 54, 55, 58, 45, 56, 54, 55, 56, 67, 79, 77, 88, 66, 89, 65, 43, 45, 54

5) Compute mean, median and mode for the following data:

Class Intervals (Marks)	Frequencies (*f*)
50- 59	12
40- 49	10
30- 39	9
20- 29	11
10- 19	8
1- 9	10
	N= 60