Measures of central tendency

Dr.K.Vyshnavi,
Assistant professor,
Dept. of community medicine,
Narayana Medical college.

Contents

- Average or central value
- Measures of central tendency
- Mean
- Median
- Mode

Average (central value)

- It is the one around which all other values are dispersed or distributed
- In most of the data, average is at the centre of concentration of values
- It carries the important properties of data
- **Central tendency**–the statistical measure that identifies a single value, as a representative of an entire distribution
- Ex: average marks of students in final MBBS
  Average number of persons visiting hospital per day

Desirable properties of a measure of central tendency:

1. should be rigidly defined
2. computation should be based on all observations
3. least affected by extreme observations
4. should be capable of further mathematical assessment

Uses:

1. To find out whether an observation is normal
2. To compare the two groups

- Mostly used to assess what is normal in health like height, birth weight, pulse rate etc.
- Normal is not just the mean or central value but accepted range of variation on either side of mean

Different measures of central tendency

- Mean (arithmetic mean)
- Median
- Mode
- Geometric mean
- Harmonic mean

Mean

- It implies arithmetic average.
- It is sum of all observations divided by number of observations.
- The most common measure of central tendency
- It is denoted by $\bar{x}$.

$$\bar{x} = \frac{\text{sum of observations}}{\text{number of observations}}$$

For ungrouped data

- Ex: weights in kgs of 10 children aged 2 years are given below calculate mean
- 9,11,10,9,12,8,9,11,10,9
- Mean$=\frac{9+11+10+9+12+8+9+11+10+9}{10}$

$=98/10=9.8$ kg

Assumed mean method:

- Used when the observations are large in size and labourious to calculate manually

Steps in calculation:

1. assume arbitrary mean/assumed mean(choose any observation from given data)
2. Subtract this value from all the given observations
3. We get differences

4.obtain mean for the differences by usual method

5. calculation of actual mean by adding the mean of differences with assumed mean

Ex: In a series of 10 postmortems following observations regarding weight of liver(gms) were found. calculate average weight of liver

- 1420,1405,1425,1410,1415,1435,1430,1415, 1445,1430
- Arbitrary mean=1420
- subtraction:(1420-1420),(1405-1420),(1425-1420)……(1430-1420)

- Differences=0,-15,5,-10,-5,15,10,-5,25,10
- Mean of differences=30/10=3
- Thus actual mean $\bar{x}$ =1420+3=1423

weight	Wt-assumed mean
1420	0
1405	-15
1425	5
1410	-10
1415	-5
1435	15
1430	10
1415	-5
1445	25
1430	10
	Sum=30

Mean for difference=30/10=3
Actual mean=1420+3=1423

For grouped data,

1. Discrete frequency distribution

- if we have $x_1, x_2, x_3 \ldots\ldots\ldots\ldots x_n$ observations with corresponding frequencies $f_1, f_2, f_3, \ldots\ldots\ldots f_n$
- Arithmetic mean is calculated as,

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + f_3x_3 + \ldots\ldots f_nx_n}{f_1 + f_2 + f_3 + \ldots\ldots f_n}$$

Ex: calculate the average number of children per family from following data.

No. of children	No. of families
0	30
1	52
2	60
3	65
4	18
5	10
6	5

No. of children(x)	No. of families(f)	Total no. of children(fx)
0	30	0
1	52	52
2	60	120
3	65	195
4	18	72
5	10	50
6	5	30
Total	240	519

$\bar{x} = 519/240 = 2.16$

2. continuous frequency distribution

In continuous frequency distribution(class interval), frequency is not associated with single value but spread over the entire class

steps in calculation,

- Write all the class intervals in first column and corresponding frequency(f) in second column
- Calculate mid value of each class interval, enter in third column(x)
- Multiply f and x, write in fourth column
- $\bar{x} = \frac{\text{sum}(fx)}{\text{sum}(f)}$

- Ex: calculate the arithmetic mean of marks obtained in biostatistics course

Marks	no. of students
40-49	4
50-59	6
60-69	16
70-79	22
80-89	9

Marks	no. of students (f)	Mid value of class interval (x)	fx
40-49	4	44.5	178
50-59	6	54.5	327
60-69	16	64.5	1032
70-79	22	74.5	1639
80-89	9	84.5	760.5
	Sum f =57		Sumfx=3936.5

$\bar{x} = 3936.5/57 = 69.06$

Summary of calculation of mean

1.Ungrouped series.
2.Grouped series.

Ungrouped series

1.Direct.
2.Indirect.

If the size of observations is small, then go for, direct method

Grouped series-discrete frequency distribution
-continuous frequency distribution

Merits and demerits

Merits :

- 1. easy to calculate and understand
- 2. it is based on all observations
- 3.most familiar to common man and rigidly defined
- 4.capable of further mathematical assessment

Demerits :

- 1.used only for quantitative data
- 2.unduly affected by extreme values

Median

- This is the next common measure after mean
- It is the value of middle observation after placing the observations in either ascending or descending order.
- Half the values lie above it and half below it.
- If n is odd, the median is the middle number.

i.e= $(n+1/2)^{th}$ observation

- If n is even, the median is the average of the two middle numbers
- Median=

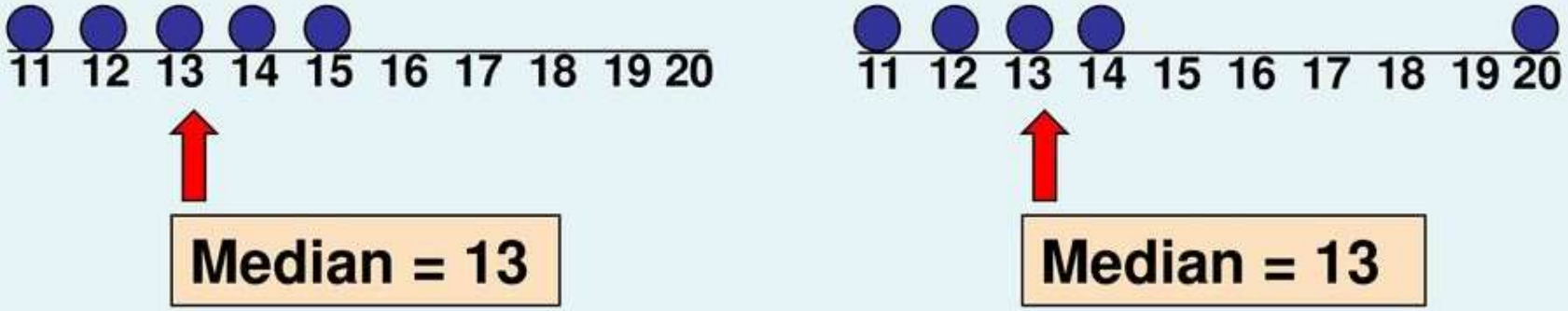
$$\frac{(\text{value of}(n/2)^{th}\text{ob} + \text{value of}(n/2+1)^{th}\text{ob})}{2}$$

Advantages:

1. can be applied for qualitative and quantitative data.
2. Not affected by extreme values

Measures of Central Tendency: The Median

- In an ordered array, the median is the “middle” number (50% above, 50% below)



- Not affected by extreme values

Ex: To find the median of 4,5,7,2,1

- Step 1 :Arrange the numbers in ascending order.
  1,2,4,5,7
- Step 2: The total elements in the distribution (5) is odd.
- The middle position can be calculated using the formula. **(n+1)/2**

  So the middle position is **(5+1)/2 = 6/2 = 3**

  The number at 3rd position is = **Median = 4**

Ex: find the median of 5,7,2,1,6,4.

step 1: arrange the numbers in ascending order.

1,2,4,5,6,7.

step 2:the total numbers in the distribution is 6

(even).

so the average of two numbers which are respectively in positions n/2 and (n/2)+1 will be the median of the given data.

Median = (4+5)/2 = 4.5.

For grouped data,(class intervals)

- 1.obtain class boundaries
- 2.calculate cumulative frequencies
- 3.calculate N/2 value=sumf/2(N=sum of freq)
- 4.locate the class interval with a cumulative frequency equal to or just more than N/2 value.(median lies in this class interval)
- 5.apply the formula

- $\text{Median} = l + \left(\frac{N/2 - cf}{f}\right) \times h$

  l=lower boundary of median class

  N=sum of frequencies

  cf=cumulative freq which is previous to median class

  f= frequency of median class

  h=width of class interval

- Ex: find the median weight of 590 infants born in a hospital, in one yr

Weight of infant in kg	No of infants
2.0–2.4	37
2.5–2.9	117
3.0–3.4	207
3.5–3.9	155
4.0–4.4	48
4.5 and above	26

Class boundaries	No of infants(f)	Cumulative freq
1.95-2.45	37	37
2.45-2.95	117	154
2.95-3.45	207	361
3.45-3.95	155	516
3.95-4.45	48	564
4.45 and above	26	590
	Sum f=N=590	

N/2=295,l=2.95,cf=154,f=207,h=0.5

Median=2.95+((295-154)/207*0.5)=3.29 kg

Mode

- The observation which occurs most frequently in a data series
- less commonly used
- Ex: Diastolic B.P values of 9 individuals is given below. Find mode
- 86,76,80,86,70,96,86,76,86
- Mode=86

- For a given data mode may not exist or it may not be unique
- Ex:5,8,6,9,2,7-no mode
- 10,18,25,9,25,9,8-2 modes

- **Geometric mean**: when data contains a few extremely large or small data or when the values change exponentially, we can use geometric mean
- For n observations it is measured as n^{th} root of product of n observations.
- May reduce to zero if any observation is zero
- **Harmonic mean**: it is the reciprocal of arithmetic mean of reciprocal observations
- HM=n/(1/x)

Selecting the appropriate measure:

Median is preferred to mean,

- Extreme scores in the distribution
- Data is measured in ordinal scale

Mode is a preferred data,

- Data is measured in nominal scale

Geometric mean is preferred

- When data are measured in logarithmic scale

conclusion

- Various measures of central tendency
- How to calculate mean for ungrouped data and grouped data
- How to calculate median
- Mode
- Selecting an appropriate measure


Thank You!

- Ex: The cholesterol levels of 10 persons are given below. calculate mean value

  260,200,240,240,260,150,220,190,210,200