Measures of Central Tendency

Arithmetic Mean

1. If the mean of 9, 8, 10, x, 12 is 20 find the value of x.

Solution:

Mean of the given numbers = (9 + 8 + 10 + x + 12)/5 = (39 + x)/5

According to the problem, mean = 15 (given).

Therefore, (39 + x)/5 = 20

⇒ 39 + x = 20 × 5

⇒ 39 + x = 100

39 - 39 + x = 100 - 39

⇒ x = 61

Hence, x = 61

Here, notice that 61 is an extreme (high) observation which affects the value of the mean drastically. This is one demerit of mean. In data which contain such extreme high or low observations , also called outliers, AM is not a very reliable measure of central tendency.

More examples on arithmetic mean:

2. If the mean of five observations x, x + 4, x + 6, x + 8 and x + 12 is 16, find the value of x.

Solution: Mean of the given observations

= x + (x + 4) + (x + 6) + (x + 8) + (x + 12)/5

= (5x + 30)/5

According to the problem, mean = 16 (given).

Therefore, (5x + 30)/5 = 16

⇒ 5x + 30 = 16 × 5

⇒ 5x + 30 = 80

⇒ 5x = 50

⇒ x = 50/5 = 10

Hence, x = 10.

Let us look at the above problem differently:

If the values are 0,4,6,8 and 12, the mean is 6. On adding a constant x to every observation gives us a mean of 6 + x. If x is 10 the mean will be 16.

3. The mean of 40 numbers was found to be 38. Later on, it was detected that a number 56 was misread as 36. Find the correct mean of given numbers.

Solution:

Calculated mean of 40 numbers = 38.

Therefore, calculated sum of these numbers = (38 × 40) = 1520.

Correct sum of these numbers

= [1520 - (wrong item) + (correct item)]

= (1520 - 36 + 56)

= 1540.

Therefore, the correct mean = 1540/40 = 38.5.

4. The mean of the heights of 6 boys is 156 cm. If the individual heights of five of them are 151 cm, 153 cm, 155 cm, 149 cm and 154cm, find the height of the sixth boy.

Solution:

Mean height of 6 boys = 156 cm.

Sum of the heights of 6 boys = (154 × 6) = 924 cm

Sum of the heights of 5 boys = (151 + 153 + 155 + 149 + 154) cm = 762 cm.

Height of the sixth boy

= (sum of the heights of 6 boys) - (sum of the heights of 5 boys)

= (924 - 762) cm = 162 cm.

Hence, the height of the sixth boy is 162 cm.

5. Find the arithmetic mean of the first 8 natural numbers.

Solution: The first 8 natural numbers are 1, 2, 3, 4, 5, 6 and 7,8

Let x denote their arithmetic mean.

Then mean = Sum of the first 8 natural numbers/number of natural numbers

x = (1 + 2 + 3 + 4 + 5 + 6 + 7 +8) / 8

= 36/8 = 4.5

If the observations are equally distributed , the mean is the average of first and last observation (1+8) /2 = 9/2 = 4.5

If there are even number of observations which are equally spaced, the mean is the middle of the centermost two observations. In this example, mean = (4+5)/2 = 4.5

If the number of equally paced observations is odd, centremost value is the arithmetic mean. Example : Calculate mean of first 7 natural numbers is 4 :

(1+2+3+4+5+6+7)/7= 28/7=4. Hence, their mean is 4.

Grouped data examples :

The following table gives the timings for 21 participants of sprint race in a sports meet. Find the average time taken .

Seconds	**Frequency**	**Midpoint x (class mark)**	**Frequency f**	**Midpoint × Frequency fx**
51 - 55	2	53	2	106
56 - 60	7	58	7	406
61 - 65	8	63	8	504
66 - 70	4	68	4	272
		Totals:	**21**	**1288**

Estimated Arithemtic Mean = *Sum of (Midpoint × Frequency) /* **Sum of Frequency**

1288/21= 61.33 seconds.

We use the word “estimated” mean because we do not know the actual time taken by each runner in the grouped data. So we assume that the frequencies are uniformly distributed over each class interval and conveniently take middle value of each class interval as x. Therefore in the first class when we say that x is 53 , we are assuming that the average time of two runners is 53 second (we do not know their actual timings). In reality, it could be possible that both runners timed 52 seconds, or that one timed 51 seconds and the other timed 54 seconds, but we make an assumption of x =53 to get our estimated mean.

It is important that each observation in the data has same unit of measurement. If one value is expressed in minutes and others in seconds, the former should also be converted to seconds.

AM will always be unique for a given data. It is well defined too. The unit of mean will always be same as the unit of measurement of the observations.

Median : It is the centremost value in a data that divides the data into 2 equal parts. Median is a positional average. A value is a median by virtue of its position.

Median in Raw Data : Arrange all observations in an ascending (or descending) order. The centremost value is the median idf there are odd number of observations. If there are evn number of observations , average of two centremost values is the median.

Example : Consider the data of 7 observations. 4, 7, 9, 10,11, 13, 14 . In this case median is the 4^{th} observation i.e. 10.

If the third observation is 8 instead of 9, the median will still remain the same. Likewise, if the sixth of seventh observations are of different value , median will not change. So median is not a good representative of all observations of the data .

Consider data with eight observations: 4, 7, 9, 10,11, 13, 14, 19. The median is (10+11)/2 = 5.5

If the 4^{th} or 5^{th} observations are changed, the median will be affected but it will not be affected if any other value is changed.

Thus **Median ={(n+1)/2}th** observation (if N is odd)

Median = (n/2th observation + (n/2 +1)th observation)/2 (If N is even)

The main advantage of the median over mean is that it is **not unduly affected by extreme values or outliers** (very high and very low values). Thus, it gives an individual a better idea of representative value. For instance, if weights of 5 people are in kg are 50, 55, 55, 60 and 150. Mean is (50+55+55+60+150)/5 = 74 kg. However, 74 kg is not a true representative value as the majority of the weights are in the 50 to 60 range. Let us calculate the median in such a case. It would be (5+1)/2th term = 3rd term. The third term is 55 kg, which is a median. Since the majority of the data is in 50 to 60 range, 55 kg is a true representative value of the data.

However n case there are no extreme observations in the data , AM is always preferred over median as it is based on all observations and more representative.

For grouped data,

M = **Estimated Median** =

$$M_m = l + \left(\frac{\frac{n}{2} - cf}{f}\right)h$$

Where

l = lower limit of median class,

n = number of observations,

cf = cumulative frequency of class preceding the median class,

f = frequency of median class,

h = class size (assuming class size to be equal)

Image credit: math.stackexchange.com

In cases, where the classes are unequal, we can calculate median and h is the class size of median class. However, AM (the above image is from be calculated for data with unequal classes because of the underlying assumption that data is uniformly distributed over the class interval, therefore we can use class mark.

For our example: Let us calculate Median

Seconds	**Frequency**	**Cumulative Frequency**
51 - 55	2	2
56 - 60	7	9
61 - 65	8	17

66 - 70	4	21

In the above table , cumulative frequency column shows that (21+1/)2= 11 th observation will exist in the class 61-65 as this contains 10th to 17th observation . Up to 9th observation is included in class 56-60. Therefore, median class is 61-65.

L=60.5 , h= 5 , cf= 9 , fm = frequency of median class = 8. Substitute in above formula and you get the answer

Median =60.5 + 0.9375

= 61.4375