



MAT2001-Statistics for Engineers

Median

Median of data is the value of the variable which divides arranged data into two equal parts. When the observations in the data are arranged either in ascending (increasing) order or in descending (decreasing) order, the Median is defined as the value of the variable, which divides the arranged data into two equal parts. We denote it by \tilde{X}

Case A : Raw data : When the number of observations in the data is an 'odd' number, then the Median is equal to the middle most term in the arranged data. If the number of observation is even, the simple average of the two middle most values of the arranged data is median.

Case B :
Discrete frequency distribution : In this case, median is defined as the value of the variable that corresponds the less than cumulative frequency just above half the total frequency.

Case C :
Continuous frequency distribution : In the case of continuous frequency distribution, median can be obtained by using the following formula :

$$\text{Median} = \tilde{X} = L + \left(\frac{\frac{N}{2} - m}{f} \right) C$$

where, L = Lower boundary of the median class

f = Frequency of the median class

m = Less than cumulative frequency of the class that precedes the median class

C = Length of the median class

N = Total Frequency

Here, Median class is the class that corresponds the less than cumulative frequency just greater than $(N/2)$ value.

Remark : When, we are given the class intervals in descending order, the median can be found by using the following formula.



MAT2001-Statistics for Engineers

$$\text{Median} = \tilde{X} = U - \left(\frac{\frac{N}{2} - m}{f} \right) C$$

where, 'U' is the upper bound of the median class. Other symbols such as N , m , f and C remain the same as in the previous formula.

USES OF MEDIAN

It is used frequently in practice. It is readily used in cases where quantitative measurement of all items is difficult but ordering of items is relatively easy. When the data contains extreme values, it is advantageous to use median. Its usefulness as a positional average is recognised in statistical analysis.

Remarks :

1. For a continuous frequency distribution Median can be found by plotting gives.
2. The sum of the absolute deviations of observations taken from the median is always minimum, i.e., $\sum |X - \tilde{X}| \leq \sum |X - A|$, where A is any arbitrary value.

SOLVED PROBLEMS ON MEDIAN

Problem 2.10.1: Determine Median for the following data : 26, 20, 15, 45, 18, 8, 10, 38, 13.

Solution : Arrange the observations in ascending order as
8, 10, 13, 15, 18, 20, 26, 38, 45
Number of observations = 9 (odd number)
Value of the middle term = Median = 18.

Problem 2.10.2: Find the median for the following values :
16, 12, 5, 8, 9, 5, 10, 28.

Solution : Arrange the observations in ascending order as,
5, 5, 8, 9, 10, 12, 16, 28.
Number of observations = 8 (even number)
Median = Average of values of two middle terms

$$\text{Median} = \frac{9+10}{2} = 9.5$$



MAT2001-Statistics for Engineers

Problem 2.10.3: Calculate the median for the following frequency distribution.

Variable X	3	4	5	6	7	8
Frequency (f)	2	4	7	9	10	16

Solution :

X	f	Less than cum. f
3	2	2
4	4	6
5	7	13
6	9	22
7 Median	10	32 Median
8	16	48
Total	N = 48	-

N/2 = 24

From the above table, $\frac{N}{2} = \frac{48}{2} = 24$

Here, '32' is just greater than $\left(\frac{N}{2}\right)$ value.

∴ Median = The value of the variable that corresponds less than cumulative frequency just greater than N/2 value = 7.

Problem: The weight of 30 students are given in the following frequency distribution. Determine the median weight of the students.

Weight in kgs	60-64	65-69	70-74	75-79	80-84
No. of students	2	8	12	5	3

Solution :

Class Boundaries	F	Less than cum. f
59.5-64.5	2	2
64.5-69.5	8	10 = m
69.5-74.5 Median class	12 = f	22
74.5-79.5	5	27
79.5-84.5	3	30
Total	N = 30	-

N/2 = 15

From table, $\frac{N}{2} = \frac{30}{2} = 15$ Here, 69.5 – 74.5 is the median class.

Dr. D. Ramesh Kumar, M.Sc., M.Phil., Ph.D.,



MAT2001-Statistics for Engineers

We have, $L = 69.5$; $m = 10$; $f = 12$; and $C = 5$

$$\begin{aligned} \text{Median} &= L + \left(\frac{\frac{N}{2} - m}{f} \right) C \\ &= 69.5 + \left[\frac{15 - 10}{12} \right] 5 = 69.5 + \frac{25}{12} = 69.5 + 2.08 = 71.58 \\ \text{Median weight} &= 71.58 \text{ kg.} \end{aligned}$$

Problem: Find the median weight for the following table:

Weight in kgs	84-80	79-75	74-70	69-65	64-60
No. of students	3	5	12	8	2

Solution: The class intervals are given in the descending order:

Class Boundaries	f	Less than cum. f
84.5-79.5	3	3
79.5-74.5	5	8 = m
74.5-69.5 Median class	12 = f	20
69.5-64.5	8	28
64.5-59.5	2	30
Total	30	-

Here 74.5 – 69.5 is the median class.

From table, $\frac{N}{2} = \frac{30}{2} = 15$

$U = 74.5$; $m = 8$, $f = 12$, and $C = 5$

$$\begin{aligned} \text{Median} &= \bar{X} = U - \left(\frac{\frac{N}{2} - m}{f} \right) C \\ &= 74.5 - \left[\frac{15 - 8}{12} \right] 5 = 74.5 - \frac{35}{12} \\ &= 74.5 - 2.92 = 71.58 \\ \text{Median weight} &= 71.58 \text{ kg.} \end{aligned}$$



MAT2001-Statistics for Engineers

Problem: In the frequency distribution of 170 families given below, the number of families corresponding to expenditure groups 20-30 and 40-50 are missing in the table. However, the median is known to be 35. Find the missing frequencies.

Expenditure	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of families	10	20	?	40	?	25	15

Solution : Let the missing frequencies of the expenditure groups 20-30 and 40-50 be f_1 and f_2 respectively.

Given , Median = 35 and total frequency = 170

Class	f	Less than cum. f
0-10	10	10
10-20	20	30
20-30	f_1	$30 + f_1 = m$
30-40 Median class	$40 = f$	$70 + f_1$
40-50	f_2	$70 + f_1 + f_2$
50-60	25	$95 + f_1 + f_2$
60-70	15	$110 + f_1 + f_2$
Total	170	-

Since, median = 35, the median class is given by 30-40.

From the table, we have

$$L = 30, f = 40, m = 30 + f_1$$

$$\frac{N}{2} = \frac{170}{2} = 85 \text{ and } C = 10$$

$$\therefore \text{Median} = L + \left(\frac{\frac{N}{2} - m}{f} \right) C$$

$$\Rightarrow 35 = 30 + \left[\frac{85 - (30 + f_1)}{40} \right] 10$$

$$\Rightarrow 35 - 30 = \frac{55 - f_1}{4}$$

$$\Rightarrow 55 - f_1 = 20 \text{ or } f_1 = 55 - 20 = 35$$

Since, the total frequency $N = 170$, or $110 + f_1 + f_2 = 170$, and $f_1 = 35$, we get $110 + 35 + f_2 = 170$ or $f_2 = 170 - 145 = 25$

\therefore The missing frequencies are respectively $f_1 = 35$ and $f_2 = 25$.