**Mode**

The mode of a set of observations is that value which occurs with the maximum frequency. This definition, properly speaking, applies to a discrete variable only.

For a continuous variable, the above definition needs to be modified. The mode here is the ***value of the variable with the highest frequency-density*** corresponding to the ideal distribution which would be obtained if the total frequency were increased indefinitely and if, at the same time, the width of the class-intervals were decreased indefinitely. Graphically, it may be looked upon as the abscissa corresponding to the highest ordinate in the frequency curve (the limiting form of histogram).

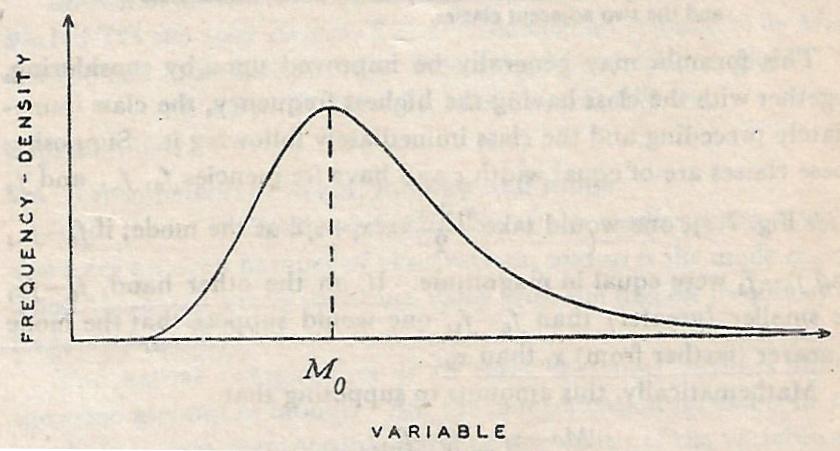


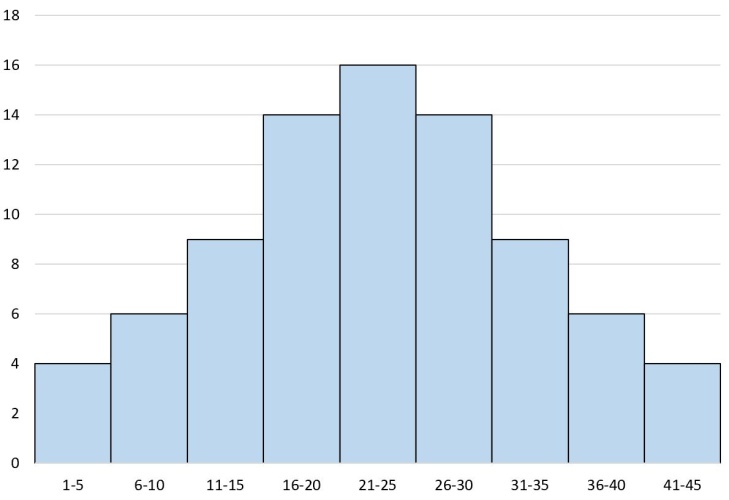
Fig. 2: Mode of a continuous variable

**Three approaches for computation of mode from grouped frequency distribution**

1. By the application of simple interpolation in a frequency distribution
2. Graphical method
3. From empirical relation

**Computation of mode by application of simple interpolation in a frequency distribution**

Suppose, the frequency distribution is perfectly symmetric as follows:



Then, **t**he mid-value of the class interval having the highest frequency (i.e. modal class) may be approximately taken to be the mode.

Thus, if the lower and upper class-boundaries of the class containing the highest frequency are and respectively, then the mode will be approximately given by

, where is the width of the interval ......................(1)

But in reality the frequency distribution will **rarely be symmetric.** Suppose, the frequency distribution is as follows:

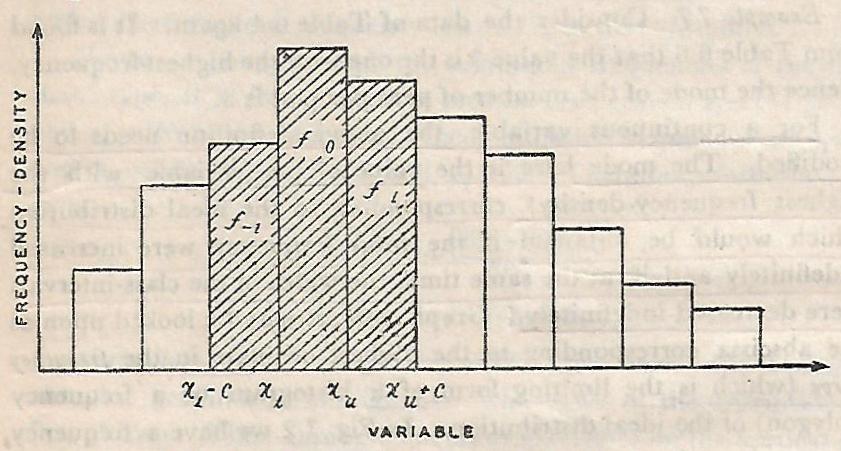


Fig. 3: Frequency distribution of a variable

In this case, one would not expect the mode to be in the middle of the modal class, i.e. . Rather, the mode will be nearer to or depending on the frequencies in the classes immediately preceding and immediately following the modal class.

Suppose the modal class, preceding class and the following class have frequencies , and respectively. Then, if (-) be smaller (greater) than (-), one would suppose that the mode is nearer (further from) than . Mathematically, this amounts to supposing that

where, is the difference of frequencies in the modal class and the preceding class and is the difference of frequencies in the modal class and the following class.

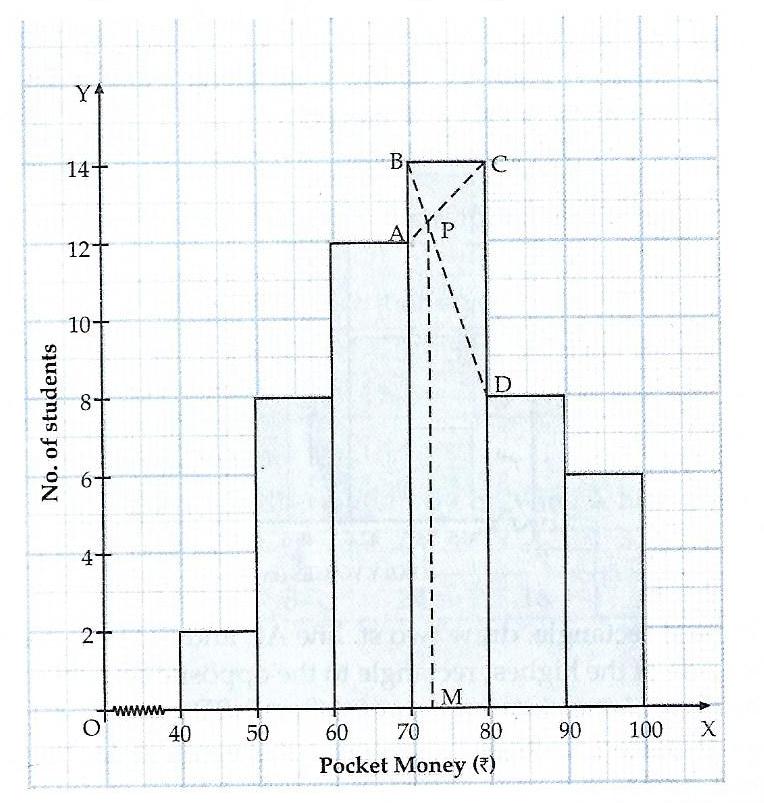
**Determination of mode using Graphical method**

The mode value may also be obtained geometrically from the histogram. Inside the highest rectangle two straight lines from the corners of the rectangles on either side of the highest rectangle to the opposite corners of the highest rectangle are drawn. The measure of mode is given by the abscissa of the point of intersection of the two straight lines.

**Example:** In a school, the weekly pocket money of 50 students is as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Weekly pocket money (in Rs.) | 40-50 | 50-60 | 60-70 | 70-80 | 80-90 | 90-100 |
| No. of students | 2 | 8 | 12 | 14 | 8 | 6 |

Draw a histogram and find the mode from the graph.



**Relation between Mean, Median and Mode**

* For unimodal distributions of moderate skewness the following approximate relation has been found to hold:
* When the distribution is symmetrical, mean, median and mode coincide, i.e.

In particular, for the ***normal distribution*** mean, median and mode are all equal.

* In most frequency distributions, it has been observed that the three measures of central tendency (mean, median and mode) obey the approximate relation provided the distribution is not very skew.

**Determination of mode using the Empirical Relation**

Another method of determining the mode is to make use of the above mentioned empirical relation, i.e. , which implies that

Mode =

Given the mean and the median, an approximate value of the mode may be obtained using the above relation.

**Advantages of mode**

* From a simple frequency distribution, mode can be obtained only by inspection. Also, for a simple series with small number of observations, mode can be determined without any calculation.
* It is unaffected by the presence of extreme values
* It can be calculated from frequency distributions with open-end classes.

**Disadvantages of mode**

* Mode has no significance unless a large number of observations are available.
* It is a peculiar measure of central tendency. When all values occur with equal frequency, there is no mode. On the other hand, if two or more values have the same maximum frequency, there is more than one mode.
* It cannot be treated algebraically
* It is quite difficult to calculate mode accurately for continuous variable
* It is less reliable and less stable as regards to sampling fluctuations

**Uses of mode:**

* Mode is most useful as a measure of central tendency ***when examining categorical data***, such as models of cars or flavours of soda, for which a mathematical average or median value based on ordering cannot be calculated.
* ***In some cases, mode is desirable than mean or median***. For example,
* A dealer of shoes will be more interested in the mode of sizes of shoes he sells than arithmetic mean or median.
* Modal wage may be considered as the representative wage of a group of workers. Modal wage is that wage which the largest number of workers receives.

**Exercise 1:** Calculate the Median and Mode from the following data:

|  |  |
| --- | --- |
| **Value** | **Frequency** |
| Less than 10 | 4 |
| Less than 20 | 16 |
| Less than 30 | 40 |
| Less than 40 | 76 |
| Less than 50 | 96 |
| Less than 60 | 112 |
| Less than 70 | 120 |
| Less than 80 | 125 |

[**Ans.** Median=36.25, Mode=34.29]

**Other measures of central tendency**

Besides arithmetic mean, median and mode, there are two other measures of central tendency which are relatively unimportant but may be appropriate to particular situations. These are ***geometric mean*** and ***harmonic mean***.

**Geometric mean (GM)**

If a variable has given values, , then its geometric mean is defined as

Thus the logarithm of the geometric mean of a variable is the arithmetic mean of its logarithm. It may be noted that *geometric mean is defined only when no observation is zero*.

Let and ( = 1, 2, 3,..., ) be the values of variables and for the individual. Then,

That is, the geometric mean of the ratios of and is the ratio of their geometric means. *Owing to this property of the geometric mean, it is sometimes preferred for averaging ratios of two variables.*

If have frequencies , then its geometric mean is defined as

, where

If be the geometric means of several groups having observations respectively, then the geometric mean of the compound group is defined as

, where

**Uses of geometric mean**

* If a series of numbers are in geometric progression, either exactly or approximately, geometric mean is the appropriate average to use.
* It is, therefore, considered as the most appropriate type in finding the average rate under compound interest, depreciation of machines and in the growth of living organism.
* Geometric mean is particularly useful in economics and business in construction of index numbers.
* Geometric mean also comes in if one wants to determine the value of a variable at the mid-point of a time interval when the variable changes over time exponentially.

**Harmonic mean (HM)**

The **simple harmonic mean** () of a variable , with given values is defined as

Thus, the reciprocal of the harmonic mean of a variable is the arithmetic mean of its reciprocal. Harmonic mean is defined only when no observation is zero.

Sometimes the variable may be in the form ‘ per unit ’, e.g. miles per hour, rupees per litre, grams per cubic foot, etc. In such cases, ***the harmonic mean would be the proper average if equal units of were considered, while the arithmetic mean would be appropriate if equal units of were considered.***

**Illustration**: Suppose a train moves equal distances, each of miles, with speeds , ,..., miles per hour. Then,

, the HM of the given speed

Suppose the train moves for equal time intervals, each of length hours, with speeds , ,..., miles per hour. Then,

, the AM of the given speed

**, where**

When all the weights are equal, the weighted harmonic mean becomes as simple harmonic mean.

**Exercise 1:** A motor car covered a distance of 50 miles four times. The first time at 50 m.p.h, the second at 20 m.p.h., the third at 40 m.p.h. and the fourth at 25 m.p.h. Calculate the average speed.

**Exercise 2:** A person bought 6 rupees worth of orange from five markets at 15p, 20p, 25p, 30p and 50p per orange respectively. What is the average price of an orange? What would be the average price, if he had purchased 20 oranges from each market?

**Exercise 3:** Three groups of observations contain 8, 7 and 5 observations. Their geometric means are 8.52, 10.12 and 7.75 respectively. Find the geometric mean of the 20 observations in the single group formed by pooling the three groups.

**Exercise 4:** (a) A man travelled 12 miles at 4 m.p.h and again 10 miles at 5 m.p.h. What was the average speed?

(b) A man travelled 12 hours at 4 m.p.h and again 10 hours at 5 m.p.h. What was the average speed?

**Exercise 5:** The arithmetic mean of two observations is 25 and their geometric mean is 15. Find (i) their harmonic mean and (ii) the two observations.

**Exercise 6:** If and are two positive values of a variate, prove that their geometric mean is equal to the geometric mean of their arithmetic and harmonic means.

[**Ans**. E1: 30 m.p.h; E2: 24p, 28p; E3: 8.837; E4: (a) 4.40 m.p.h, (b) 4.45 m.p.h; E5: (i) 9, (ii) 45, 5]

**Quantiles or Partition values**

Just as median divides the total number of observations into two equal parts, there are similar other measures *which are used to divide or partition the observations into a fixed number of parts say* ***4, 10*** *or* ***100****.* The general name for a measure of this type is ***quantiles*** or **fractiles** or partition values. ***A frequency distribution may be briefly described by giving the values of some of its quantiles.***

Some of the important types of quantiles or partition values are 1) Median, 2) Quartiles, 3) Deciles and 4) Percentiles.

**Median:** It is the middle most value of a set of observations, i.e. it divides the total number of observations into two equal parts. The number of observations smaller than median is the same as the number larger than it. For data of continuous type, exactly one-half of the observations are smaller than median, i.e. median is the value of the variable corresponding to cumulative frequency . These ideas are extended to Quartiles, Deciles and Percentiles.

**Quartiles:** Quartiles are such values (of the variable) which divide the total number of observations into 4 equal parts. Obviously, there are 3 quartiles-

1. First quartiles (or Lower quartile):
2. Second quartile (or Middle quartile):
3. Third quartile (or Upper quartile:

The number of observations smaller than is the same as the number lying between and or between and Q3 or larger than . For data of continuous type, one-quarter of the observations is smaller than , two-quarters are smaller than and three-quarters are smaller than . This implies that , and are values of the variable corresponding to less-than type cumulative frequencies , and respectively. Since =, it is evident that the second quartile is the same as median.

Measures of central tendency, dispersion and skewness can be obtained based on quartiles. For example,

* (a measure of central tendency)
* (a measure of dispersion)

**Some uses of quartiles**

Quartiles often are used in sales and survey data to divide populations into groups. For example, one can use quartiles to find the top 25 percent of incomes in a population. Some companies use the quartiles to benchmark other companies.

In the real world often range are used to represent the amount of spread in the data. For example, temperature ranges for the day on a weather report, min/max levels of water in a reservoir. In the presence of outliers, IQR is a better representation of the amount of spread in the data rather than the range. This is because in computation of IQR the bottom 25% of the data points and the top 25% of the data points are ignored and thus IQR statistic is more robust with respect to outliers.

Computation of the standard deviation is extremely difficult or impossible when the observations are given in a frequency table with class-intervals of varying width or with one or both of the terminal classes undefined. Then quartile deviation may be used to represent the measure of dispersion.

**Deciles:** Deciles are such values which divide the total number of observations into 10 equal parts. There are 9 deciles ,,..., corresponding to cumulative frequencies ,,..., respectively. .

Deciles are used in [finance](https://cleartax.in/g/terms/finance) and [economics](https://cleartax.in/g/terms/economics) to divide data into sets for the purpose of analysis. For example, the data set of mutual fund portfolio or the data set of income-tax return filers can be divided for analysing the top 10 per cent and so on.

**Percentiles:** Percentiles are such values which divide the total number of observations into 100 equal parts. There are 99 percentiles ,,...,, called first percentile, second percentile and so on. The percentile () is, therefore, that value of the variable up to which lie exactly of the total number of observations. Hence corresponds to less-than type cumulative frequency . In other words, the -percentile is value of the variable such that a proportion of the total number of given values are less than or equal to it and a proportion (1-) are greater than or equal to it.

***Percentiles are a great tool to use when one need to know the relative standing of a value.*** Where does a value fall within a distribution of values? It is commonly used to report scores in competitive examinations. For example, the 70th percentile on the GRE was 156. That means if one scored 156 on the exam, his score was better than 70 percent of test takers.

**Example:** Marks obtained by 200 students in examination are given below.

|  |  |
| --- | --- |
| **Marks** | **No. of students** |
| 0-10 | 5 |
| 10-20 | 10 |
| 20-30 | 14 |
| 30-40 | 21 |
| 40-50 | 25 |
| 50-60 | 34 |
| 60-70 | 36 |
| 70-80 | 27 |
| 80-90 | 16 |
| 90-100 | 12 |

Draw an ogive for the given distribution. From the graph, find

1. The median
2. The upper quartile
3. The number of students scoring above 65 marks
4. If 10 students qualify for merit scholarship, find the minimum marks required to qualify

