**Statistical descriptors**

**Numerical descriptors of data**

•The numerical methods utilize a set of characteristic numbers to summarize the data and highlight their main features.

•The most important purpose of these descriptive numerical summaries is for statistical inference, a role that graphs cannot fulfill.

**Measures of central tendency**

A central tendency refers to a central value or a representative value of a statistical series.

An average is a single value which represents the whole set of figures and all other individual items concentrate around it.

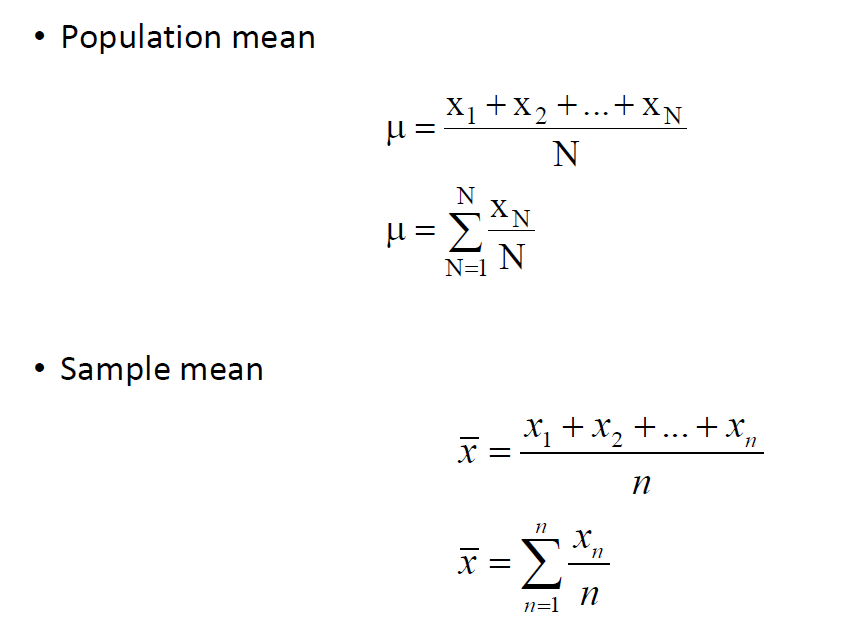
In other words, an average is a single value within the range of the data that is used to represent all the values in the series. Such an average is somewhere within the range of the data, it is therefore called as measure of central tendency or average value or measure of location.

•Generally data from many natural systems as well as from those devised by people, tend to cluster aroundsome values of variables.

•A particular value, known as the central value, can be taken as a representative of the sample.

•This feature is called central tendency because the spread seems to take place about a center.

* Mean
* Median
* Mode

**Mean**

Mean is the balance point of data, when each point is stacked on a number line.

It is the most commonly used measure of central tendency.

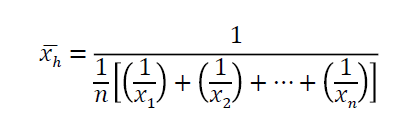
* It is easy to calculate and easy to understand
* Takes into account all available information (of data) i.e. It should be based on all observations.
* It should be uniquely defined.
* It should be capable of further algebraic treatment.
* The only real disadvantage is that it may be adversely affected by extreme values, that is, it is not resistant (Robust) to changes in the presence of a few outlying observations.

**Other means**

•There are two other measures of central tendency that are used in special situations.

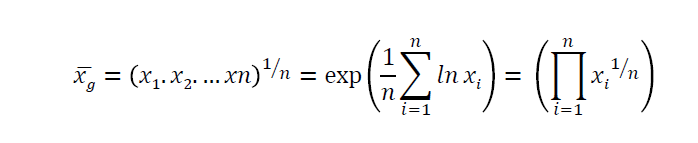
•**Harmonic mean**

The harmonic mean is the reciprocal of the mean of the reciprocals. It is applied in situations where the reciprocal of a variable is averaged.



•**Geometric mean**

The geometric mean is used in averaging values that represent a rate of change. Here the variable follows an exponential, that is, a logarithmic law. For a sample of observations, x1, x2,..., xn, the geometric mean is the positive nth root of the product of the n values. This is the same as the antilog of the mean of the logarithms



**Median**

The median is the middle number of the series when ordered.

Median = X [(n+1)/2], when n is odd

And

Median = A.M of X(n/2) and X[(n/2)+1], when n is even

•When a summary value is desired that is not strongly influenced by a few extreme observations, the median is preferable to mean.

**Note**

Mean may be pulled towards the outlier. This is not true in case of median, which is a resistant measure of central tendency.

**Mode**

The mode of a set of observations is that value which occurs most often or with the greatest frequency.

**Note**

•The mode does not always exist. This is certainly true when all observations occur with the same frequency.

•For some sets of data there may be several values occurring with the greatest frequency in which case we have more than one mode.

•Determination of mode does not require any calculation and it can be used for qualitative as well as quantitative data.

•Thus, if jogging is the preferred exercise by most people, we say that jogging is the modal choice.