Skewness

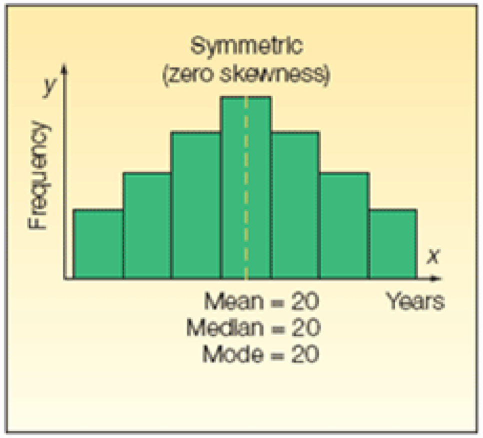
A distribution in which the values equidistant from the mean have equal frequencies and is called **Symmetric Distribution**.

Any departure from symmetry is called **skewness**.

In a perfectly **symmetric distribution, Mean=Median=Mode** and the two tails of the distribution are equal in length from the mean. These values are pulled apart when the distribution departs from symmetry and consequently one tail become longer than the other.

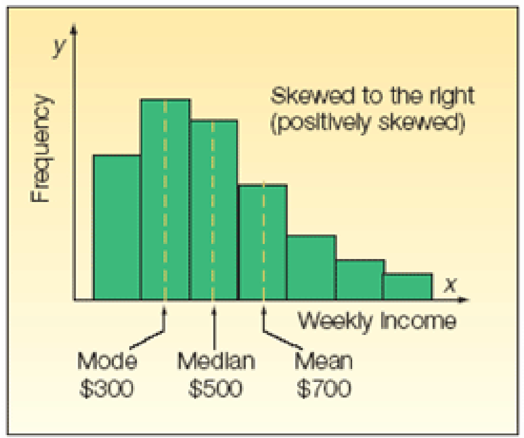
If right tail is longer than the left tail then the distribution is said to have **positive skewness**. In this case, **Mean>Median>Mode**

If left tail is longer than the right tail then the distribution is said to have **negative skewness**. In this case, **Mean<Median<Mode**



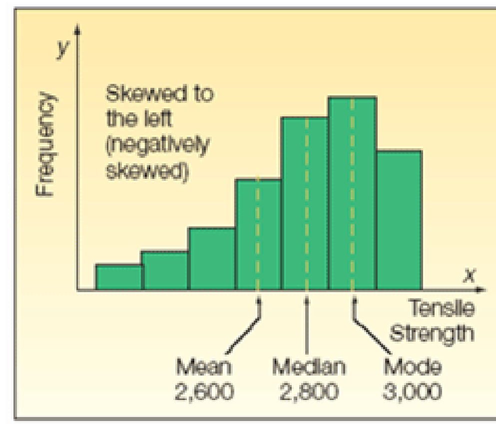
zero skewness

mode = median = mean



positive skewness

mode < median < mean



negative skewness

mode > median > mean

Skewness

When the distribution is symmetric, the value of skewness should be zero.

Karl Pearson defined coefficient of Skewness as:

$$Sk = \frac{Mean - Mode}{SD}$$

Since in some cases, Mode doesn't exist, so using empirical relation,

$$Mode = 3Median - 2Mean$$

We can write,

$$Sk = \frac{3\left(Median - Mean\right)}{SD}$$

(it ranges b/w -3 to +3)

Skewness

According to Bowley (a British Statistician):

Bowley's coefficient of skewness (also called Quartile skewness coefficient)

$$sk = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_1 - 2Q_2 + Q_3}{Q_3 - Q_1} = \frac{Q_1 - 2Median + Q_3}{Q_3 - Q_1}$$

Another measure of skewness mostly used is by using moment ratio (denoted by $\sqrt{\beta_1}$):

$$sk = \frac{\mu_3}{\sigma^3}, \qquad \text{for population data}$$

$$sk = \frac{m_3}{s^3}, \qquad \text{for sample data}$$

For symmetric distributions, it is zero and has positive value for positively skewed distribution and take negative value for negatively skewed distributions.

Kurtosis

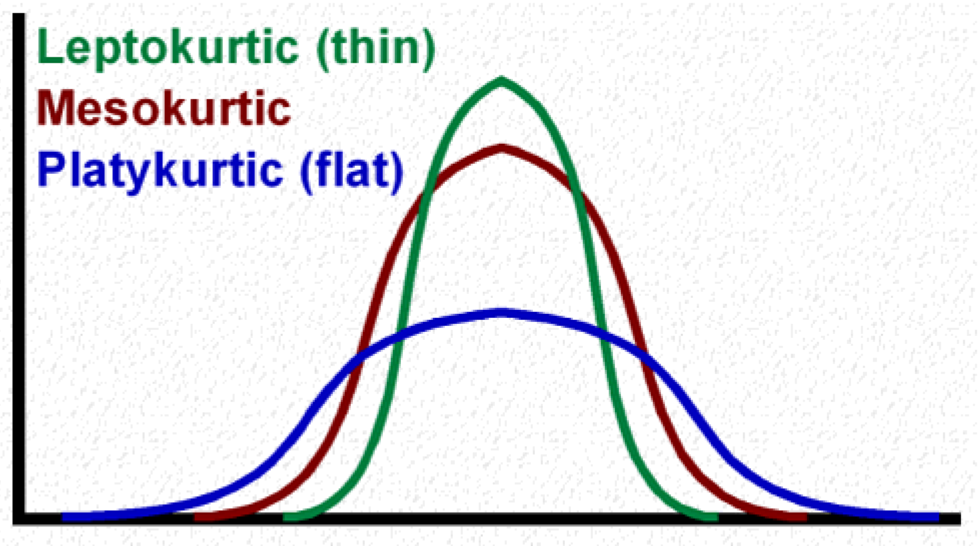
Karl Pearson introduced the term Kurtosis (literally the amount of hump) for the degree of peakedness or flatness of a unimodal frequency curve.

When the peak of a curve becomes relatively high then that curve is called Leptokurtic.

When the curve is flat-topped, then it is called Platykurtic.

Since normal curve is neither very peaked nor very flat topped, so it is taken as a basis for comparison.

The normal curve is called Mesokurtic.



Kurtosis

Kurtosis is usually measured by the moment ratio (β_2).

$$Kurt = \beta_2 = \frac{\mu_4}{\mu_2^2}, \qquad \text{for population data}$$

$$Kurt = b_2 = \frac{m_4}{m_2^2}, \qquad \text{for sample data}$$

For a normal distribution, kurtosis is equal to 3.

When is greater than 3, the curve is more sharply peaked and has narrower tails than the normal curve and is said to be leptokurtic.

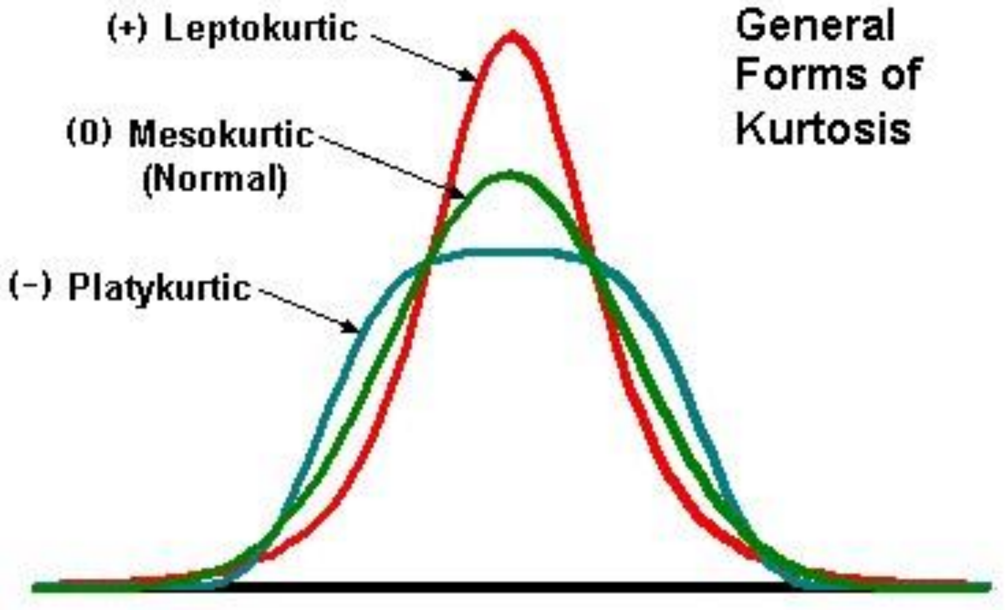
When it is less than 3, the curve has a flatter top and relatively wider tails than the normal curve and is said to be platykurtic.

Kurtosis

Excess Kurtosis (EK): It is defined as:

EK=Kurtosis-3

- For a normal distribution, EK=0.
- When EK>0, then the curve is said to be Leptokurtic.
- When EK<0, then the curve is said to be Platykurtic.



Kurtosis

Another measure of Kurtosis, known as Percentile coefficient of kurtosis is:

$$\text{Kurt}=\frac{Q.D}{P_{90}-P_{10}}$$

Where,

Q.D is semi-interquartile range=Q.D=$(Q_3-Q_1)/2$

P_{90}=90th percentile

P_{10}=10th percentile

You will recall that the Pearson's coefficient of skewness is defined as

(mean - mode)/standard deviation,

and if we apply the empirical relation between the mean, median and the mode, then the coefficient is given by:

Pearson’s coefficient of skewness:

$$SK = \frac{(\text{mean} - \text{median})}{\text{standard deviation}}$$

$$SK = \frac{3\,(\text{mean} - \text{median})}{\text{standard deviation}}$$

As you can see, this coefficient involves the calculation of the mean as well as the standard deviation.

Actually, the numerator is divided by the standard deviation in order to obtain a pure number.

If the analysis of a data-set is being undertaken using the median and quartiles alone, then we use a measure called Bowley's coefficient of skewness.

The advantage of this particular formula is that it requires NO KNOWLEDGE of the MEAN or STANDARD DEVIATION.

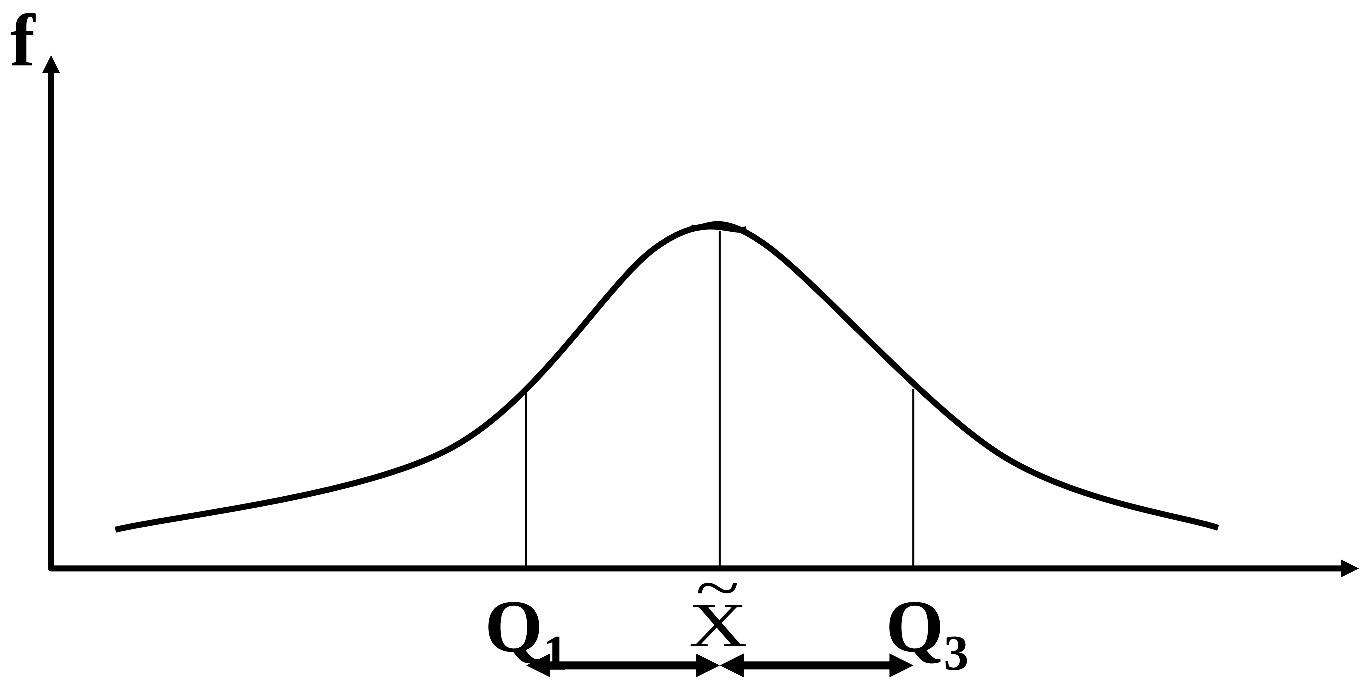
In an asymmetrical distribution, the quartiles will NOT be equidistant from the median, and the AMOUNT by which each one deviates will give an indication of skewness.

Where the distribution is positively skewed, Q_1 will be closer to the median than Q_3.

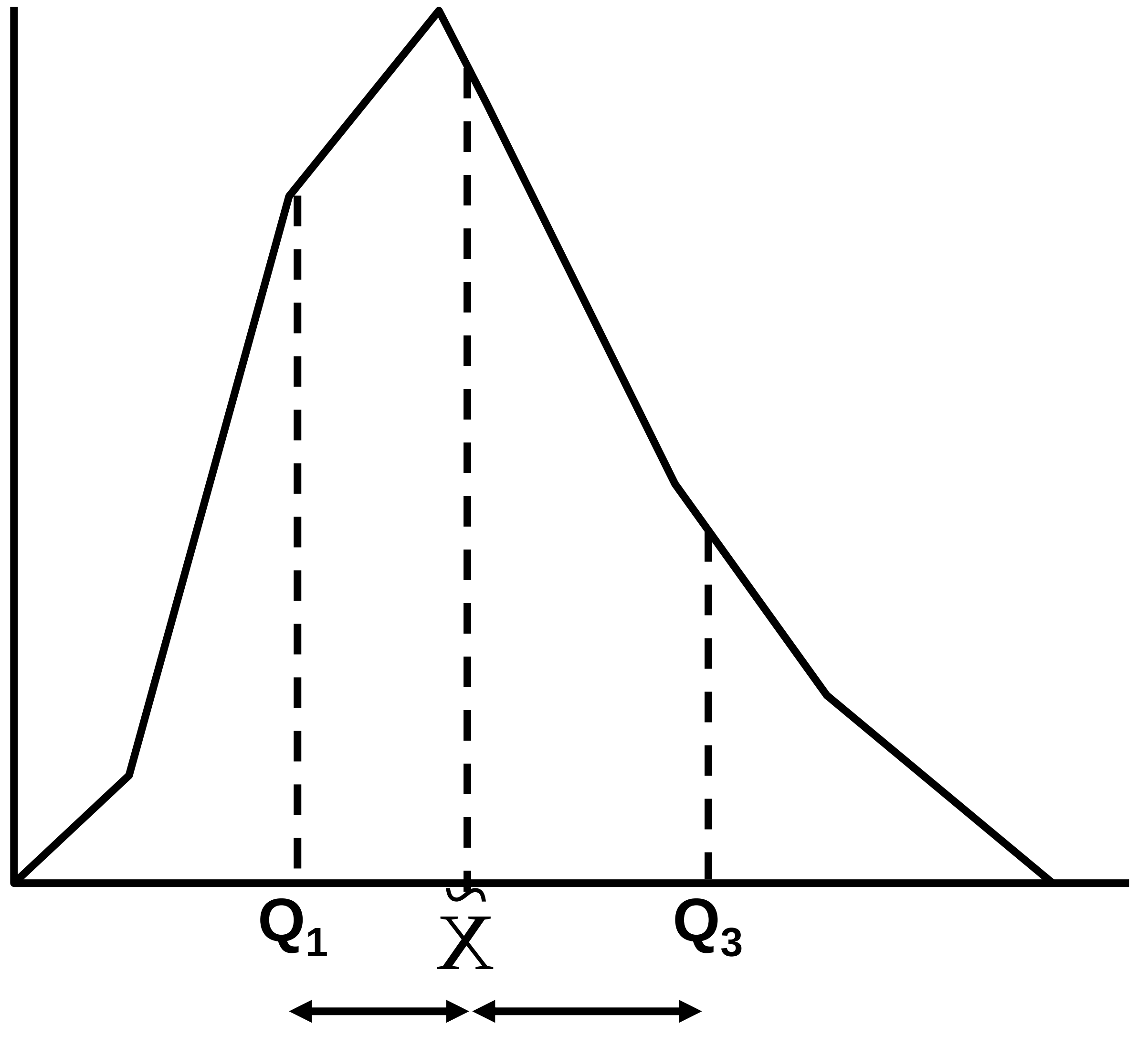
In other words, the distance between Q_3 and the median will be greater than the distance between the median and Q_1.

If the data were perfectly symmetrical, the following would be true:

1.The distance from Q_1 to the median would be equal to the distance from the median to Q_3, as shown below:



POSITIVE SKEWNESS



And hence,

if we subtract the distance median - Q_1 from the distance Q_3 - median, we will obtain a positive answer.

In case of a positively skewed distribution:

$(Q_3 - \text{median}) - (\text{Median} - Q_1) > 0$

i.e. $Q_1 + Q_3 - 2\ \text{median} > 0$

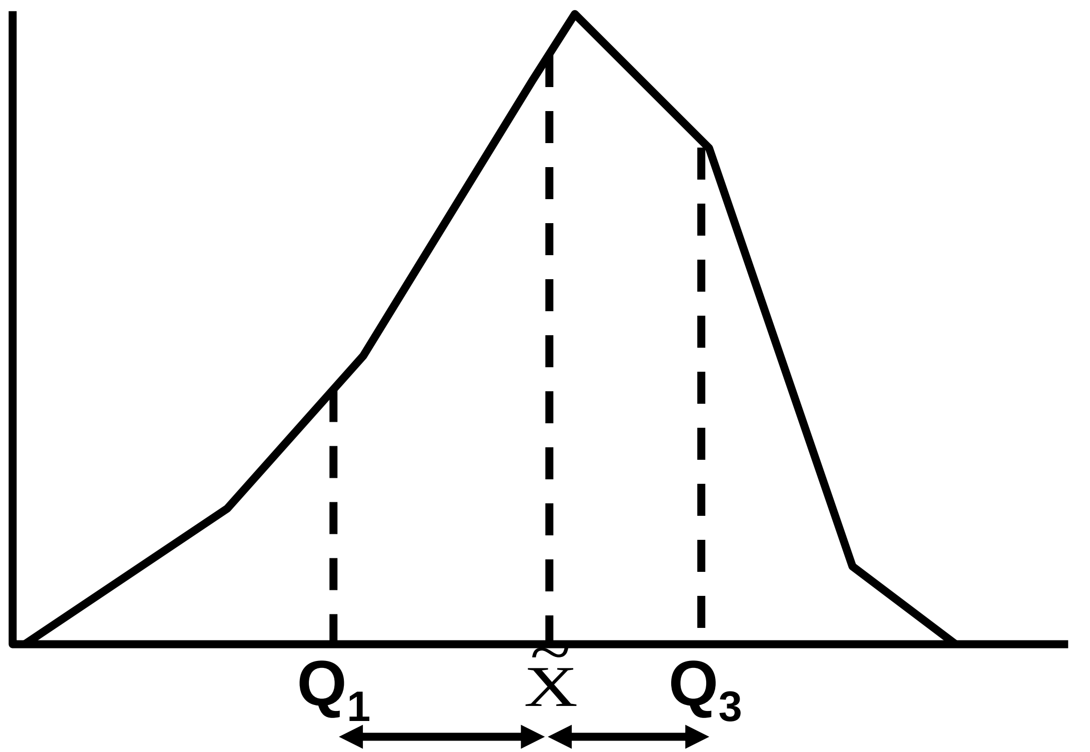
The opposite is true for skewness to the left:

NEGATIVE SKEWNESS

In this case:

$(Q_3 - \text{median}) - (\text{Median} - Q_1) < 0$

i.e. $Q_1 + Q_3 - 2\ \text{median} < 0$



By inspecting these, it can be seen that one distribution is symmetrical while the other is quite different.

The distinguishing feature here is the degree of asymmetry or *SKEWNESS* in the two polygons. In order to *measure* the skewness in our distribution, we compute the PEARSON's COEFFICIENT OF SKEWNESS which is defined as:

$$\frac{\text{mean} - \text{mode}}{\text{standard deviation}}$$

Applying the *empirical relation* between the mean, median and the mode, the Pearson's Coefficient of Skewness is given by:

$$= \frac{3\left(\text{mean} - \text{median}\right)}{\text{s tan dard deviation}}$$

For a *symmetrical* distribution the coefficient will always be ZERO, for a distribution skewed to the RIGHT the answer will always be positive, and for one skewed to the LEFT the answer will always be negative.

Let us now calculate this coefficient for the example of the children of the manual and non-manual workers. Sample statistics pertaining to the ages of these children are as follows:

	Children of Manual Workers	Children of Non-Manual Workers
Mean	8.50 years	8.50 years
Standard deviation	3.61 years	3.61 years
Median	8.50 years	9.16 years
Q_1	6.00 years	5.50 years
Q_3	11.00 years	10.83 years
Quartile deviation	2.50 years	2.66 years

The Pearson's Coefficient of Skewness is calculated for each of the two categories of children, as shown below:

Ages of Children of Manual Workers	Ages of Children of Non-Manual Workers
$\dfrac{3(8.50-8.50)}{3.61}$	$\dfrac{3(8.50-9.16)}{3.61}$
$= 0$	$= -0.55$

For the data pertaining to children of manual workers, the coefficient is zero, whereas, for the children of non-manual workers, the coefficient has turned out to be a negative number. This indicates that the distribution of the ages of the children of the manual workers is symmetric whereas the distribution of the ages of the children of the non-manual workers is negatively skewed.

The students are encouraged to draw the frequency polygon and the frequency curve for each of the two distributions, and to compare the results that have just been obtained with the *shapes* of the two distributions.

The gist of the above discussion is that in case of a positively skewed distribution, the quantity

$$Q_1 + Q_3 - 2\tilde{X}$$

will be positive, whereas in case of a negatively distribution, this quantity will be negative.

A <u>RELATIVE</u> measure of skewness is obtained by dividing $Q_1 + Q_3 - 2\tilde{X}$ by the inter-quartile range i.e. $Q_3 - Q_1$, <u>so that</u> Bowley's coefficient of skewness is given by:

Bowley's coefficient of skewness

$$= \frac{\left(Q_1 + Q_3 - 2\tilde{X}\right)}{Q_3 - Q_1}$$

It is a pure (unit less) number, and its value lies between 0 and ± 1.

For a positively skewed distribution, this coefficient will turn out to be positive, and for a negatively skewed distribution this coefficient will come out to be negative.

Let us apply this concept to the example regarding the ages of children of the manual and non-manual workers that we considered in the last lecture.

Example:

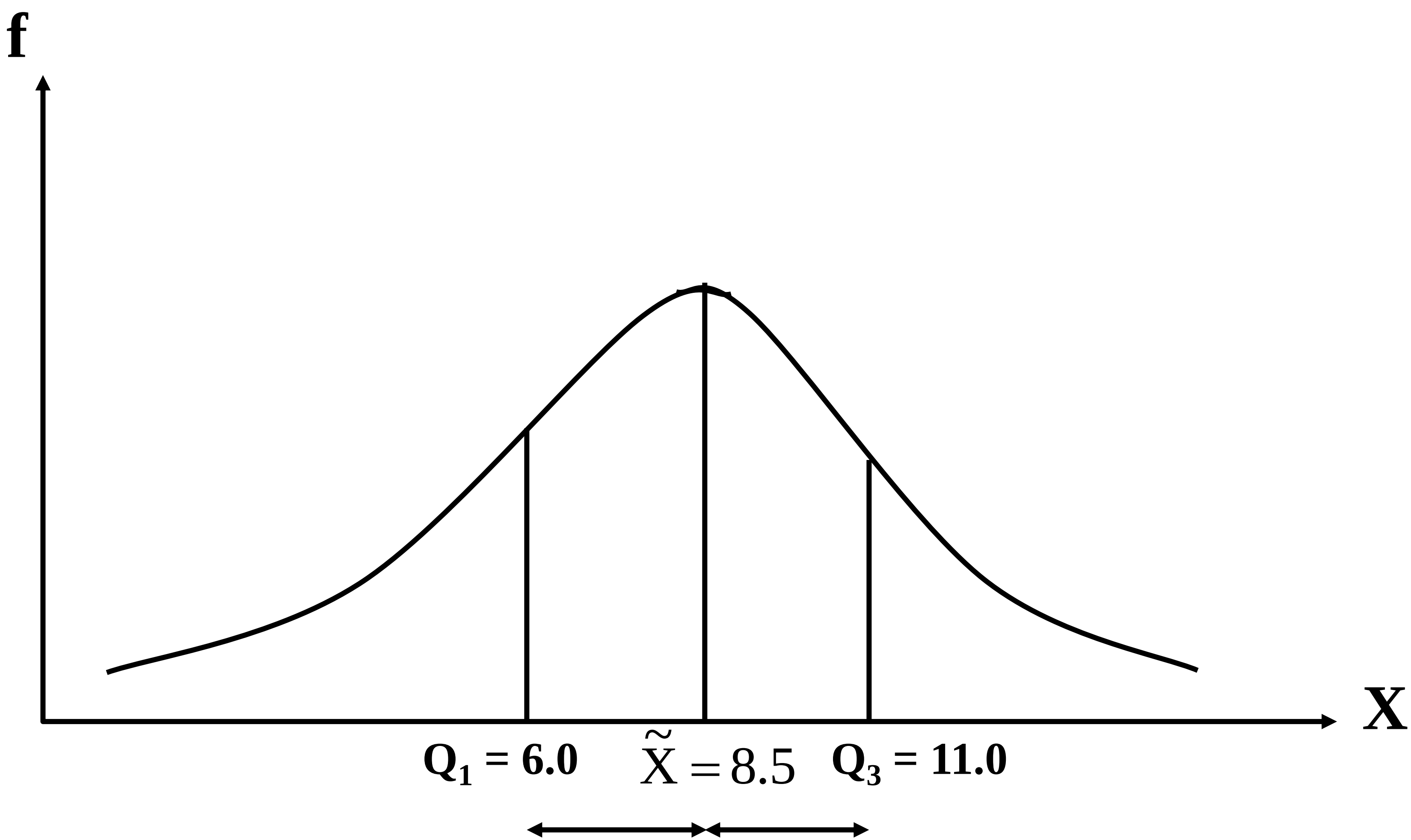
Age of Onset of Nervous Asthma in Children (to Nearest Year)	Children of Manual Workers	Children of Non-Manual Workers
0 – 2	3	3
3 – 5	9	12
6 – 8	18	9
9 – 11	18	27
12 – 14	9	6
15 – 17	3	3
	60	60

EXAMPLE:

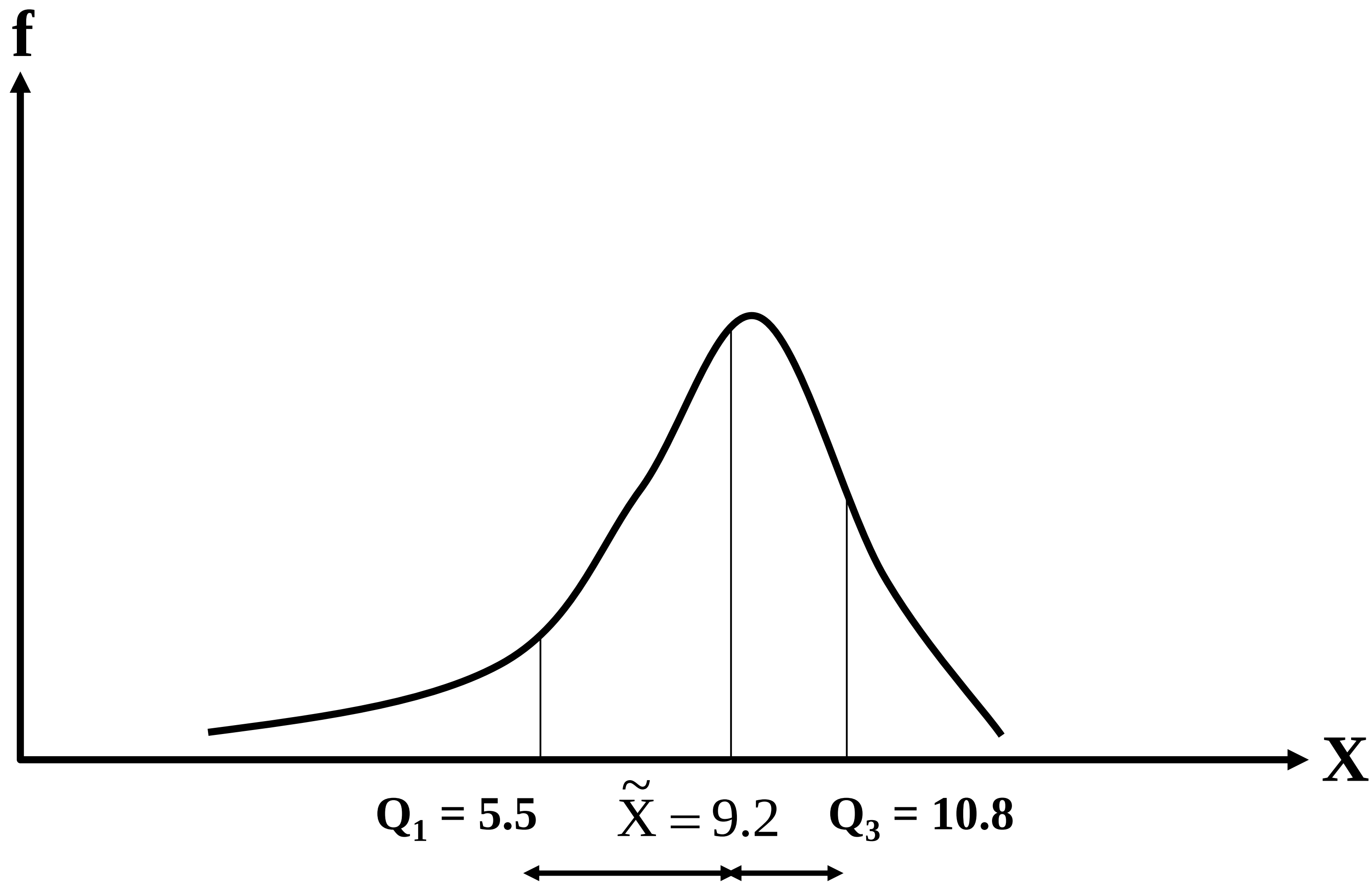
Sample statistics pertaining to ages of children of manual and non-manual workers:

	Children of Manual Workers	Children of Non-Manual Workers
Mean	8.50 years	8.50 years
Standard deviation	3.61 years	3.61 years
Median	8.50 years	9.16 years
Q_1	6.00 years	5.50 years
Q_3	11.00 years	10.83 years
Quartile deviation	2.50 years	2.66 years

The statistics pertaining to children of manual workers yield the following PICTURE:



On the <u>other</u> hand, the statistics pertaining to children of <u>non-manual</u> workers yield the following PICTURE:



The diagram pertaining to children of non-manual workers clearly shows that the distance between

$$Q_1 \text{ and } \tilde{X}$$

is much greater than the distance between

which happens whenever we are dealing with a negatively skewed distribution.

$$\tilde{X} \text{ and } Q_3$$

If we compute the Bowley's coefficient of skewness for each of these two data-sets, we obtain:

Ages of Children of Manual Workers

$$= \frac{11.00 + 6.00 - 2 \times 8.50}{2.50}$$

= 0

Ages of Children of Non-Manual Workers

$$\frac{10.83 + 5.50 - 2 \times 9.16}{10.83 - 5.50}$$

= – 0.37

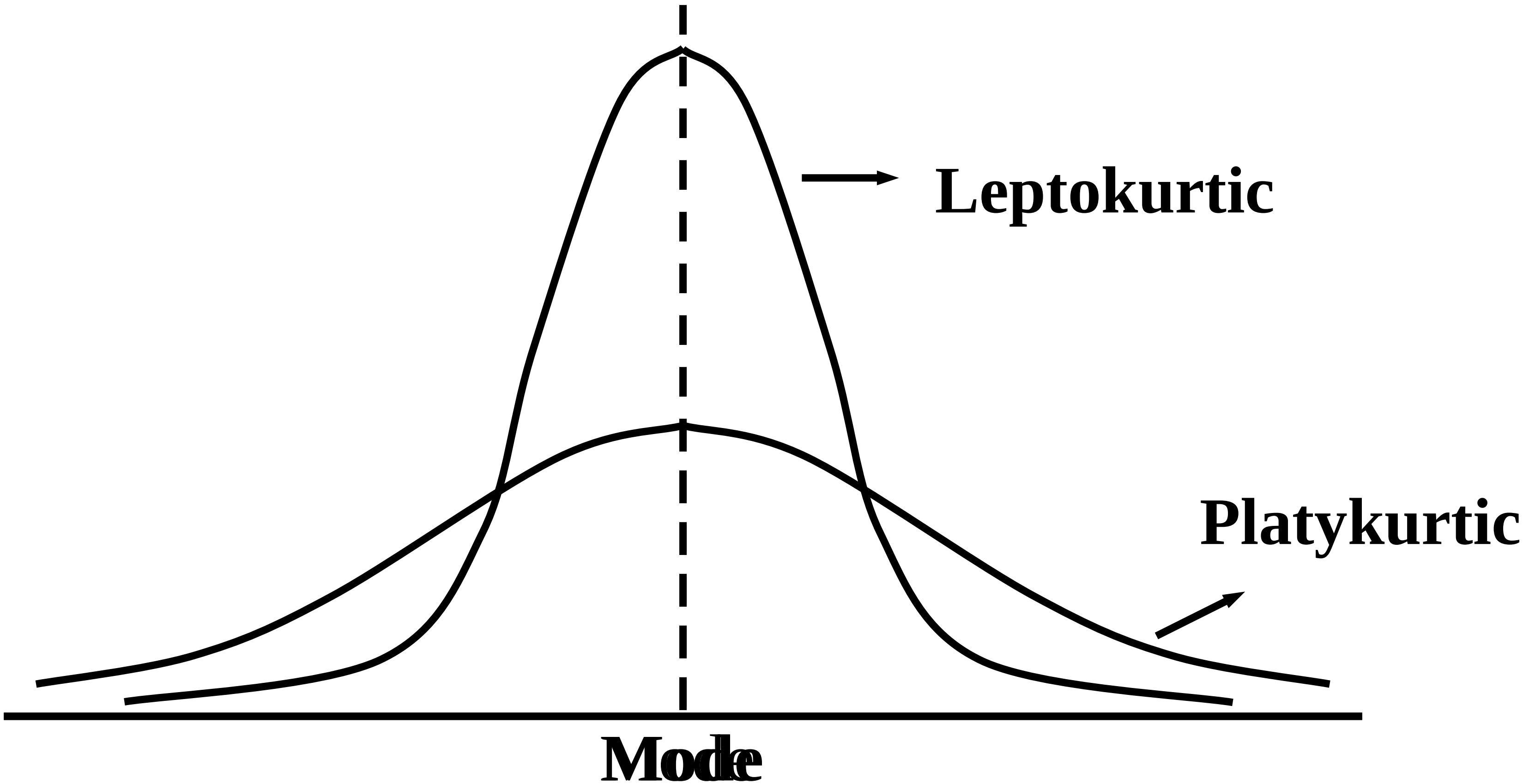
As you have noticed, for the children of the manual workers, the Bowley's coefficient has come out to be zero, whereas for the children of the non-manual workers, the coefficient has come out to be negative.

This indicates that the distribution of the ages of the children of manual workers is symmetrical whereas the distribution of the ages of the children of the non-manual workers IS negatively skewed --- EXACTLY the same conclusion that we obtained when we computed the Pearson's coefficient of skewness.

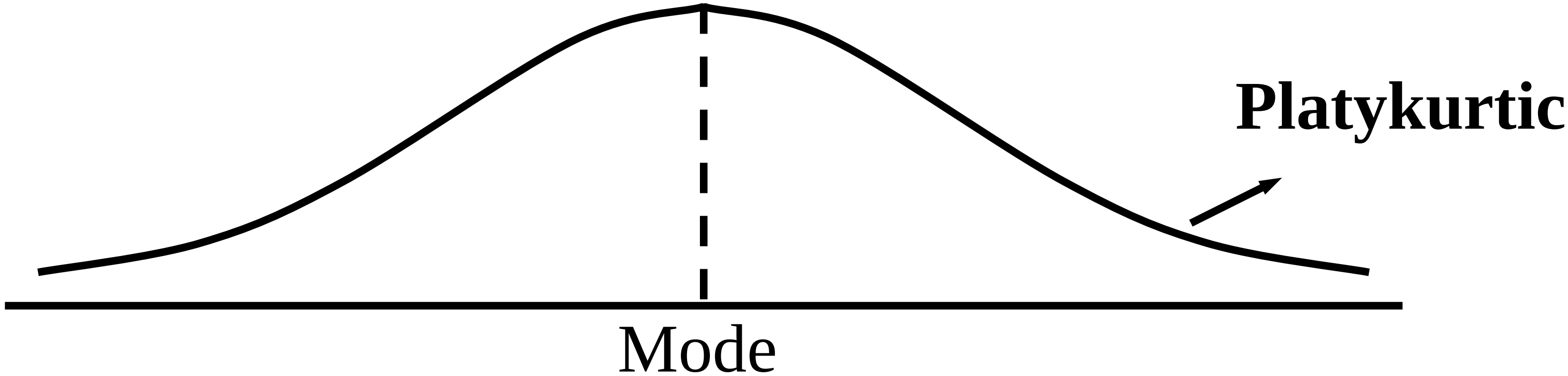
The next concept that I am going to pick-up is the concept of KURTOSIS.

The term kurtosis was introduced by Karl Pearson. This word literally means 'the amount of hump', and is used to represent the degree of PEAKEDNESS or flatness of a unimodal frequency curve.

When the values of a variable are closely BUNCHED round the mode in such a way that the peak of the curve becomes relatively high, we say that the curve is LEPTOKURTIC.


Leptokurtic
Platykurtic
Mode

On the other hand, if the curve is flat-topped, we say that the curve is PLATYKURTIC:

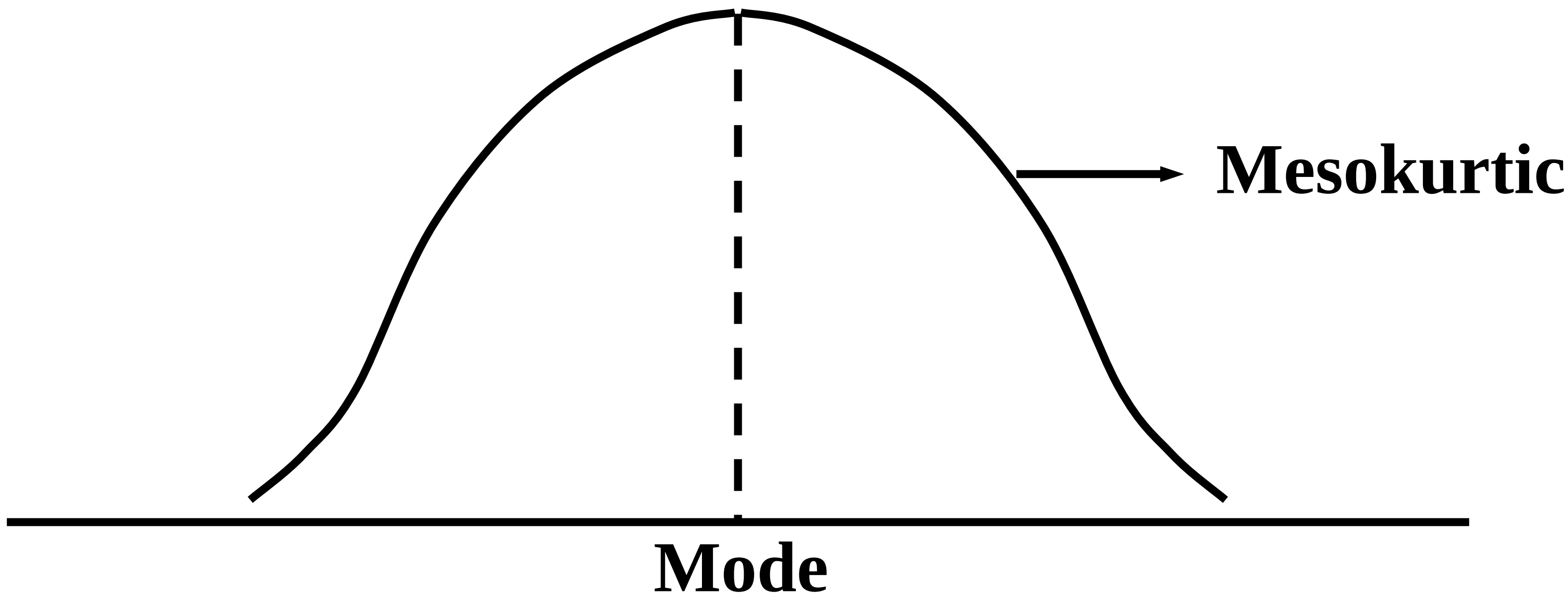

Platykurtic
Mode

The NORMAL curve is a curve which is neither very peaked nor very flat, and hence it is taken as A BASIS FOR COMPARISON.

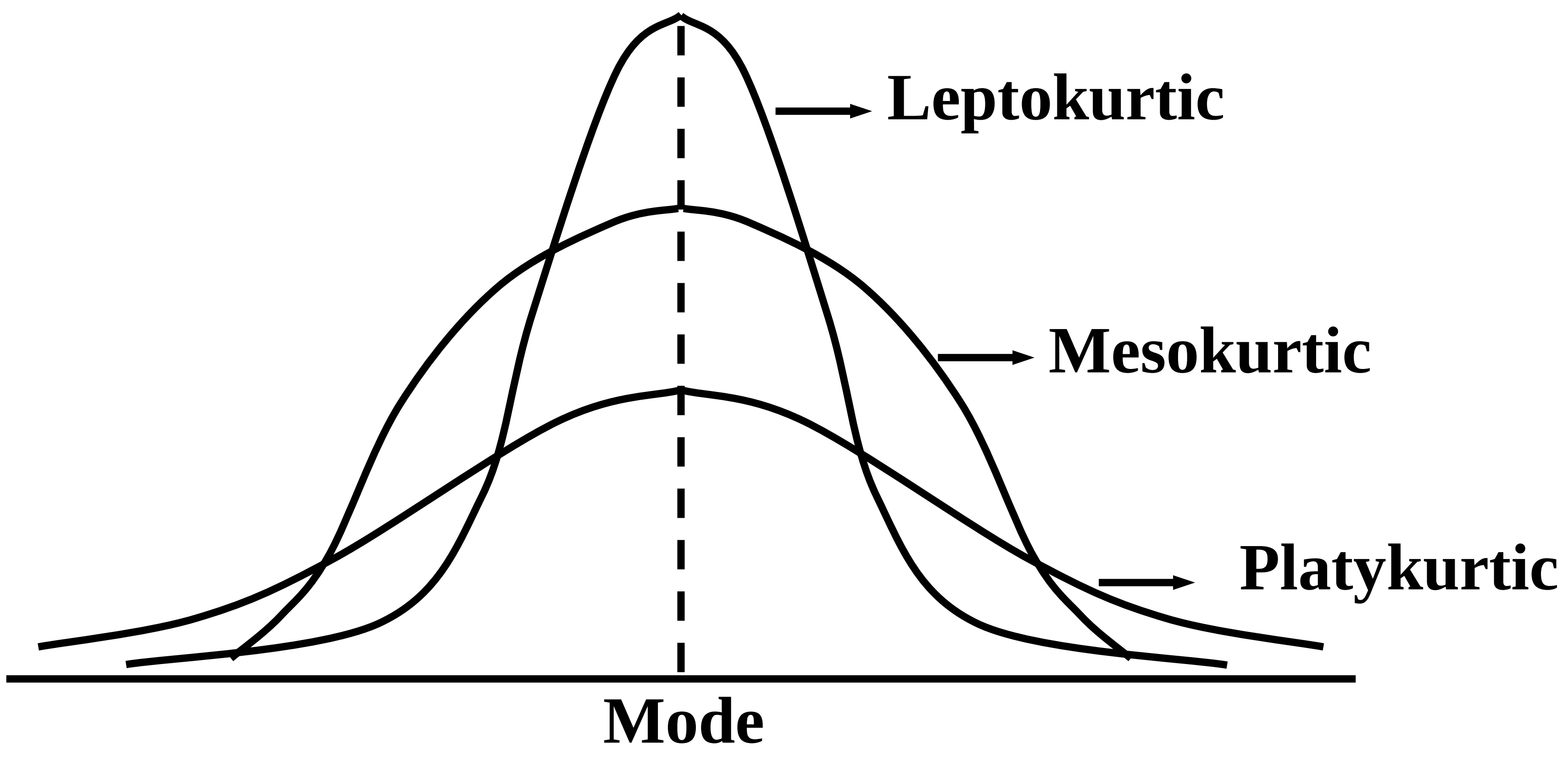
The normal curve itself is called MESOKURTIC.

I will discuss with you the normal in detail when we discuss continuous probability distributions.

At the moment, just think of the symmetric hump shaped curve shown below:


Mesokurtic
Mode

Super-imposing the three curves on the same graph, we obtain the following picture:


Leptokurtic
Mesokurtic
Platykurtic
Mode

The tallest one is called leptokurtic, the intermediate one is called mesokurtic, and the flat one is called platykurtic.

The question arises, "How will we <u>MEASURE</u> the degree of peakedness or kurtosis of a data-set?"

A MEASURE of kurtosis based on quartiles and percentiles is

$$K = \frac{Q.D.}{P_{90} - P_{10}},$$

This is known as the PERCENTILE COEFFICIENT OF KURTOSIS.

It has been shown that K for a normal distribution is 0.263 and that it lies between 0 and 0.50.

In case of a <u>leptokurtic</u> distribution, the percentile coefficient of kurtosis comes out to be <u>LESS THAN</u> 0.263,

and

in the case of a <u>platykurtic</u> distribution, the percentile coefficient of kurtosis comes out to be <u>GREATER THAN</u> 0.263.

The next concept that I am going to discuss with you is the concept of moments --- a MATHEMATICAL concept, and a very important concept in statistics.

I will request you to concentrate on this concept CAREFULLY.

MOMENTS

A *moment* designates the power to which deviations are raised before averaging them.

For example, the quantity

$$\frac{1}{n}\sum\left(x_i - \bar{x}\right)^1 = \frac{1}{n}\sum\left(x_i - \bar{x}\right)$$

is called the first sample moment about the mean, and is denoted by m_1.

Similarly, the quantity

$$\frac{1}{n}\sum\left(x_i - \bar{x}\right)^2$$

is called the second sample moment about the mean, and is denoted by m_2.

In general, the rth moment about the mean is:

the arithmetic mean of the *rth* power of the deviations of the observations from the mean. In symbols, this means that

$$m_r = \frac{1}{n}\sum\left(x_i - \bar{x}\right)^r$$

for sample data.

In a similar way, moments about an arbitrary origin, say α, are defined by the relation

, for sample data.

$$m'_r = \frac{1}{n}\sum\left(x_i - \alpha\right)^r$$

Moments about the mean are also called the <u>central moments</u> or the <u>mean moments</u>.

$$m_1' = \frac{1}{n}\sum\left(x_i - \alpha\right) = \frac{\sum x_i}{n} - \alpha = \bar{x} - \alpha.$$

For r = 1, we have

$$m_1 = \frac{1}{n}\sum\left(x_i - \overline{x}\right) = \frac{\sum x_i}{n} - \overline{x} = \overline{x} - \overline{x} = 0,$$

and

$$m_1' = \frac{1}{n}\sum\left(x_i - \alpha\right) = \frac{\sum x_i}{n} - \alpha = \overline{x} - \alpha.$$

Putting r = 2 in the relation for mean moments, we see that

$$m_2 = \frac{1}{n}\sum\left(x_i - \overline{x}\right)^2$$

which is exactly the same as the sample <u>variance</u>.

If we take the positive square root of this quantity, we obtain the <u>standard deviation</u>.

In the formula,

$$m'_r = \frac{1}{n}\sum (x_i - \alpha)^r$$

if we put $\alpha = 0$, we obtain

$$m'_r = \frac{1}{n}\sum x_i{}^r$$

and this is called the <u>rth moment about zero</u>, or the <u>rth moment about the origin</u>.

EXAMPLE:

Calculate the first four moments about the mean for the following set of examination marks: 45, 32, 37, 46, 39, 36, 41, 48 & 36.

For convenience, the observed values are written in an increasing sequence. The necessary calculations appear in the table below:

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
32	– 8	64	– 512	4096
36	– 4	16	– 64	256
36	– 4	16	– 64	256
37	– 3	9	– 27	81
39	– 1	1	– 1	1
41	1	1	1	1
45	5	25	125	625
46	6	36	216	1296
48	8	64	512	4096
360	0	232	186	10708

Now $\bar{x} = \frac{\sum x_i}{n} = \frac{360}{9} = 40$ **marks.**

Therefore

$$m_1 = \frac{\sum (x_i - \bar{x})}{n} = 0$$

$$m_2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{232}{9} = 25.78\ (\text{marks})^2$$

$$m_3 = \frac{\sum (x_i - \bar{x})^3}{n} = \frac{186}{9} = 20.67\ (\text{marks})^3$$

$$m_4 = \frac{\sum (x_i - \bar{x})^4}{n} = \frac{10708}{9} = 1189.78\ (\text{marks})^4$$

All the formulae that I have discussed until now pertain to the case of raw data.

How will we compute the various moments in the case of grouped data?

MOMENTS IN THE CASE OF GROUPED DATA:

When the sample data are grouped into a frequency distribution having k classes with midpoints x_1, x_2, …, x_k and the corresponding frequencies f_1, f_2, …,f_k ($\sum f_i = n$), the rth sample moments are given by

$$m_r = \frac{1}{n}\sum f_i\left(x_i - \bar{x}\right)^r \text{, and}$$

$$m'_r = \frac{1}{n}\sum f_i\left(x_i - \alpha\right)^r.$$

In the calculation of moments from a grouped frequency distribution, an error is introduced by the assumption that the frequencies associated with a class are located at the MIDPOINT of the class interval. You remember the concept of grouping error that I discussed with you in an earlier lecture?

Our moments therefore need corrections.

Sheppard's Corrections for Grouping Error:

It has been shown by W.F. Sheppard that, if the frequency distribution (i) is <u>continuous</u> and (ii) <u>tails off to zero at each end</u>, the <u>corrected</u> moments are as given below:

$$m_2 \text{ (corrected)} = m_2 \text{ (uncorrected)} - \frac{h^2}{12};$$

$$m_3 \text{ (corrected)} = m_3 \text{ (uncorrected)};$$

$$m_4 \text{ (corrected)} = m_4 \text{ (uncorrected)} - \frac{h^2}{2} \cdot m_2 \text{ (uncorrected)} + \frac{7}{240} \cdot h^4;$$

where h denotes the uniform class-interval.

The important point to note here is that these corrections are NOT applicable to highly skewed distributions and distributions having unequal class-intervals.

I am now going to discuss with you certain mathematical RELATIONSHIPS that exist between the moments about the mean and the moments about an arbitrary origin.

These corrections were introduced by W.F. Sheppard, and hence they are known as SHEPPARD'S CORRECTIONS.

The reason for doing so is that, in many situations, it is easier to calculate the moments in the first instance, about an arbitrary origin. They are then transformed to the mean-moments using the relationships that I am now going to convey to you.

The equations are:

$$m_1 = 0$$

$$m_2 = m'_2 - (m'_1)^2;$$

$$m_3 = m'_3 - 3\, m'_2\, m'_1 + 2(m'_1)^3, \text{ and}$$

$$m_4 = m'_4 - 4\, m'_3\, m'_1 + 6m'_2\,(m'_1)^2 - 3(m'_1)^4$$

But I would like to give you two <u>tips</u> for remembering these formulae:

1) in each of these relations, <u>the sum of the coefficients of various terms on the right hand side equals zero</u>

and

2) <u>each term on the right is of the same dimension as the term on the left</u>.

In this course, I will <u>not</u> be discussing the mathematical <u>derivation</u> of these relationships.

You are welcome to study the mathematics behind these formulae if you are interested. (The derivation is available in your own text book.)

Let us now apply these concepts to an example:

Compute the first four moments for the following distribution of marks after applying Sheppard's corrections:

Marks out of 20	5	6	7	8	9	10	11	12	13	14	15
No. of Students	1	2	5	10	20	51	22	11	5	3	1

If we wish to compute the first four moments about the mean by the direct method, first of all, we will have to compute mean itself.

The mean of this particular data-set comes out to be 10.06.

But, 10.06 is not a very convenient number to work with!

This is so because when we construct the columns of etc., $X-\overline{X},\ \left(X-\overline{X}\right)^2$ **we will have a lot many decimals.**

An alternative way of computing the moments is to take a convenient number as the arbitrary origin and to compute the moments about this number. Later, we utilize the relationships between the moments about the mean and the moments about the arbitrary origin in order to find the moments about the mean.

In this example, we may select 10 as the arbitrary origin, which is the X-value corresponding to the highest frequency 51, and construct the column of D which is the same as X-10. Next, we compute the columns of fD, fD^2, fD^3, and so on.

Earnings in Rs.(x_i)	No. of Men f_i	D_i $(x_i - 10)$	f_iD_i	$f_iD_i^2$	$f_iD_i^3$	$f_iD_i^4$
5	1	– 5	– 5	25	– 125	625
6	2	– 4	– 8	32	– 128	512
7	5	– 3	– 15	45	– 135	405
8	10	– 2	– 20	40	– 80	160
9	20	– 1	– 20	20	– 20	20
10	51	0	0	0	0	0
11	22	1	22	22	22	22
12	11	2	22	44	88	176
13	5	3	15	45	135	405
14	3	4	12	48	192	768
15	1	5	5	25	125	625
Sum	131	..	8	346	74	3718
Sum ÷ n	1	..	0.06 $=m'_1$	2.64 $=m'_2$	0.56 $=m'_3$	28.38 $=m'_4$

Moments about the mean are:

$m_1 = 0$

$m_2 = m'_2 - (m'_1)^2 = 2.64 - (0.06)^2 = 2.64$

$m_3 = m'_3 - 3m'_2 m'_1 + 2(m'_1)^3$

$= 0.56 - 3(2.64)(0.06) + 2(0.06)^3$

$= 0.08$

$m_4 = m'_4 - 4m'_3 m'_1 + 6m'_2 (m'_1)^2 - 3(m'_1)^4$

$= 28.38 - 4.(0.56)(0.06) + 6(2.64)(0.06)^2 - 3(0.06)^4 = 28.30$

Applying Sheppard's corrections, we have

$$m_2 \text{ (corrected)} = m_2 \text{ (uncorrected)} - \frac{h^2}{12}$$

$$= 2.64 - 0.08 = 2.56,$$

$$m_3 \text{ (corrected)} = m_3 \text{ (uncorrected)} = 0.08,$$

$$m_4 \text{ (corrected)} = m_4 \text{ (uncorrected)} - \frac{h^2}{2} \cdot m_2 \text{ (uncorrected)} + \frac{7h^4}{240}$$

$$= 28.30 - 1.32 + 0.03 = 27.01$$

I have discussed with you in <u>quite</u> a lot of detail the concept of moments.

The question arises, “<u>Why</u> is it that we are going through all these lengthy calculations? What is the <u>significance</u> of computing moments?”

You will obtain the answer to this question when I discuss with you the concept of <u>moment ratios</u>.

There are certain ratios in which <u>both</u> the numerators and the denominators are <u>moments</u>. The most <u>common</u> of these moment-ratios are denoted by b_1 and b_2, and defined by the relations:

Moment Ratios:

$$b_1 = \frac{(m_3)^2}{(m_2)^3} \text{ and } b_2 = \frac{m_4}{(m_2)^2}$$

(in the case of sample data)

They are independent of origin and units of measurement, i.e. they are <u>pure</u> numbers.

b_1 is used to measure the <u>skewness</u> of our distribution, and b_2 is used to measure the <u>kurtosis</u> of the distribution.

<u>First</u> let us consider b_1:

INTERPRETATION OF b_1:

- For symmetrical distributions, b1 is equal to zero. Hence, for any data-set, b1 comes out to be zero, we can conclude that our distribution is symmetric.
- Noted that the measure which will indicate the direction of skewness is the third moment round the mean.
- If our distribution is positively skewed, m_3 will be positive, and if our distribution is negatively skewed, m_3 will be negative.
- b_1 will turn out to be positive in both situations because it is given by

$$b_1 = \frac{(m_3)^2}{(m_2)^3}$$

(Since m_3 is being squared, b_1 will be positive regardless of the sign of m_3.)

INTERPRETATION OF b_2:

For the normal distribution, $b_2 = 3$.

For a leptokurtic distribution, $b_2 > 3$, and for a platykurtic distribution, $b_2 < 3$.

You have noted that the third and fourth moments about the mean provide information about the skewness and the kurtosis of our data-set. This is so because m_3 occurs in the numerator of b_1 and m_4 occurs in the numerator of b_2.

What about the dispersion and the centre of our data-set?

Do you not remember that the second moment about the mean is exactly the same thing as the variance, the positive square root of which is the standard deviation --- the most <u>important</u> measure of dispersion?

And, what about the centre of the distribution?

You will be interested to note that the first moment about zero is <u>NONE OTHER</u> than the arithmetic mean!

This is so because

$$\frac{1}{n}\sum\left(x_i - 0\right)^1$$

is equal to

$$\frac{1}{n}\sum x_i$$

--- <u>none</u> other than the arithmetic mean!

In <u>this</u> way,

the first four moments play a <u>KEY</u> <u>role</u> in describing frequency distributions.

IN TODAY'S LECTURE,

- Bowley's coefficient of skewness
- The Concept of Kurtosis
- Percentile Coefficient of Kurtosis
- Moments & Moment Ratios
- Sheppard's Corrections
- The Role of Moments in Describing Frequency Distributions

IN THE NEXT LECTURE, YOU WILL LEARN

- **The Concept of <u>Simple Linear Regression and Correlation</u>**

- **A <u>Brief</u> Introduction Multiple Linear Regression and Correlation**

Describing a Frequency Distribution

To describe the major characteristics of a frequency distribution, we need to calculate the following five quantities:

- The total number of observations in the data.
- A measure of central tendency (e.g. mean, median etc.) that provides the information about the center or average value.
- A measure of dispersion (e.g. variance, SD etc.) that indicates the spread of the data.
- A measure of skewness that shows lack of symmetry in frequency distribution.
- A measure of kurtosis that gives information about its peakeness.

Describing a Frequency Distribution

It is interesting to note that all these quantities can be derived from the first four moments.

For example,

- The first moment about zero is the arithmetic mean
- The second moment about mean is the variance.
- The third standardized moment is a measure of skewness.
- The fourth standardized moment is used to measure kurtosis.

Thus first four moments play a key role in describing frequency distributions.