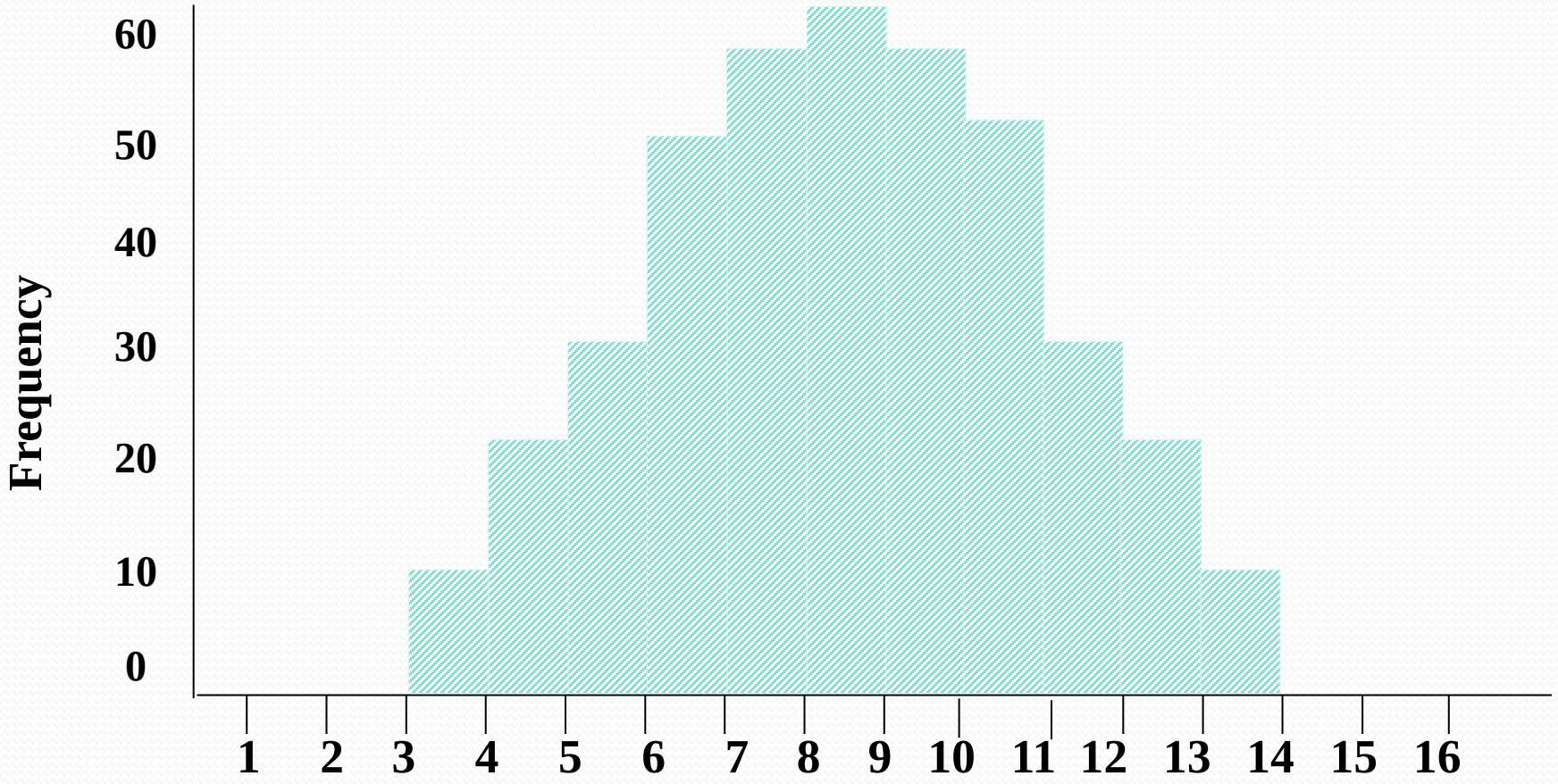

بسم الله الرحمن الرحيم
In the name of ALLAH
the most Beneficent and the most merciful

NORMAL DISTRIBUTION BY DR. ABDUL RAUF

NORMAL DISTRIBUTION

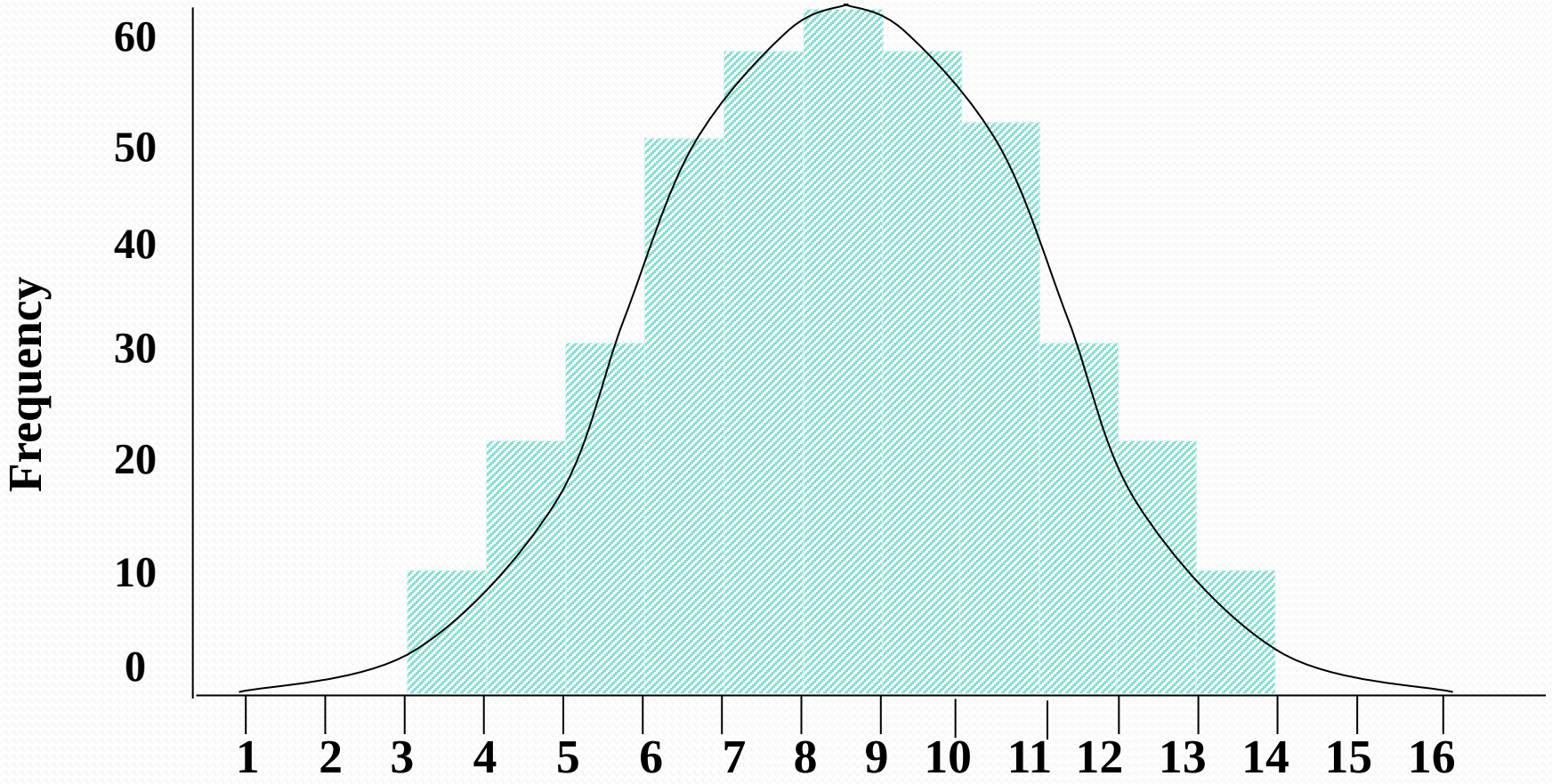
- The normal distribution or 'normal curve' is an important concept in statistical theory. Let us suppose, we collect the haemoglobin values of a very large number of people and make a frequency distribution with narrow class intervals, we are likely to get a smooth, symmetrical curve. Such a curve is called a normal distribution or normal curve.

Histogram



Normal curve

Normal Curve is a smooth symmetrical bell shaped curve as shown below.



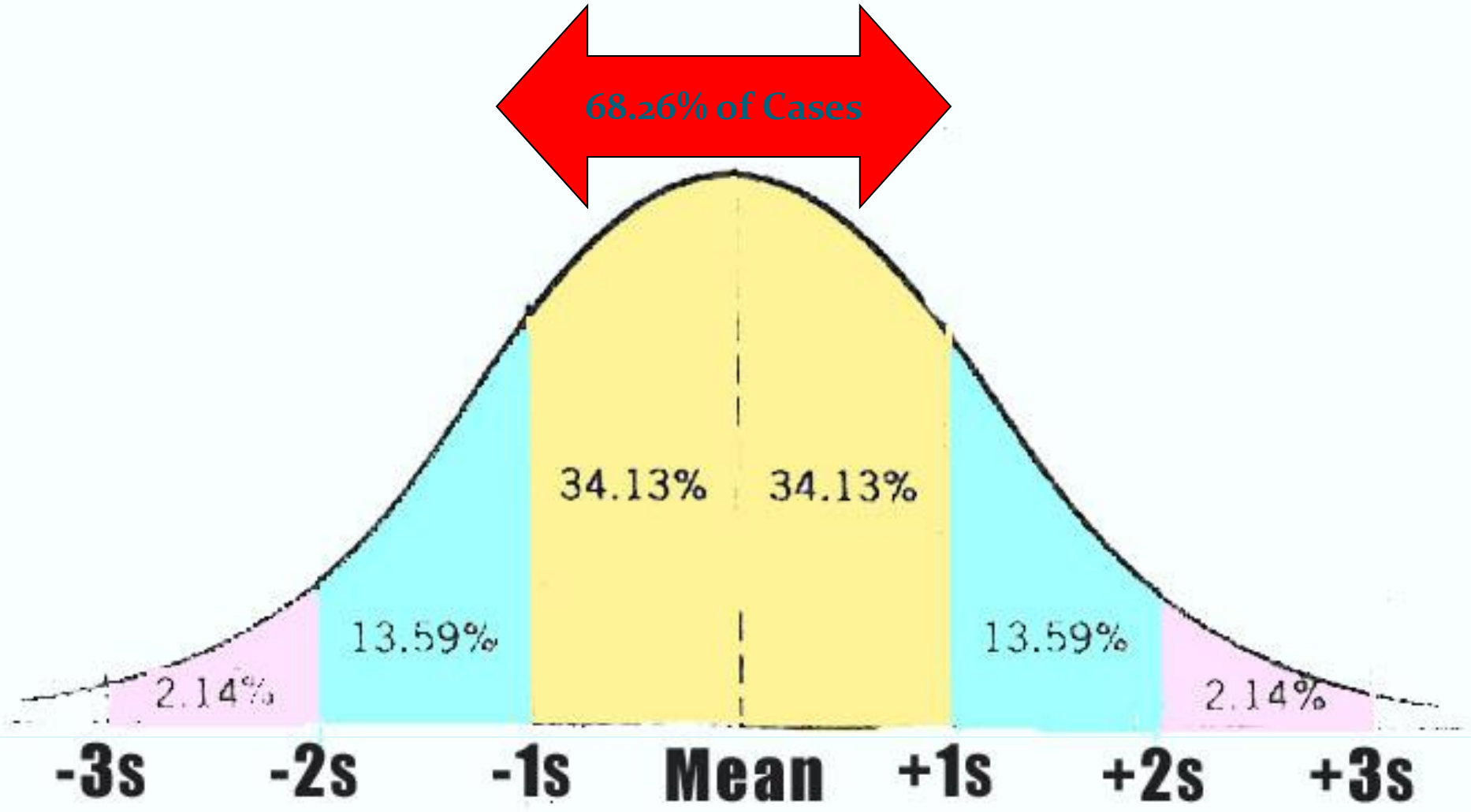
NORMAL DISTRIBUTION

- Such a curve is called a normal distribution or normal curve. The shape of the curve will depend upon the mean and standard deviation which in turn will depend upon the number and nature of observations. It follows, therefore, there will be an infinite number of normal curves.

NORMAL DISTRIBUTION

- (a) The area between one standard deviation on either side of the mean ($\bar{x} \pm 1\ \sigma$) will include approximately 68 per cent of the values in the distribution.

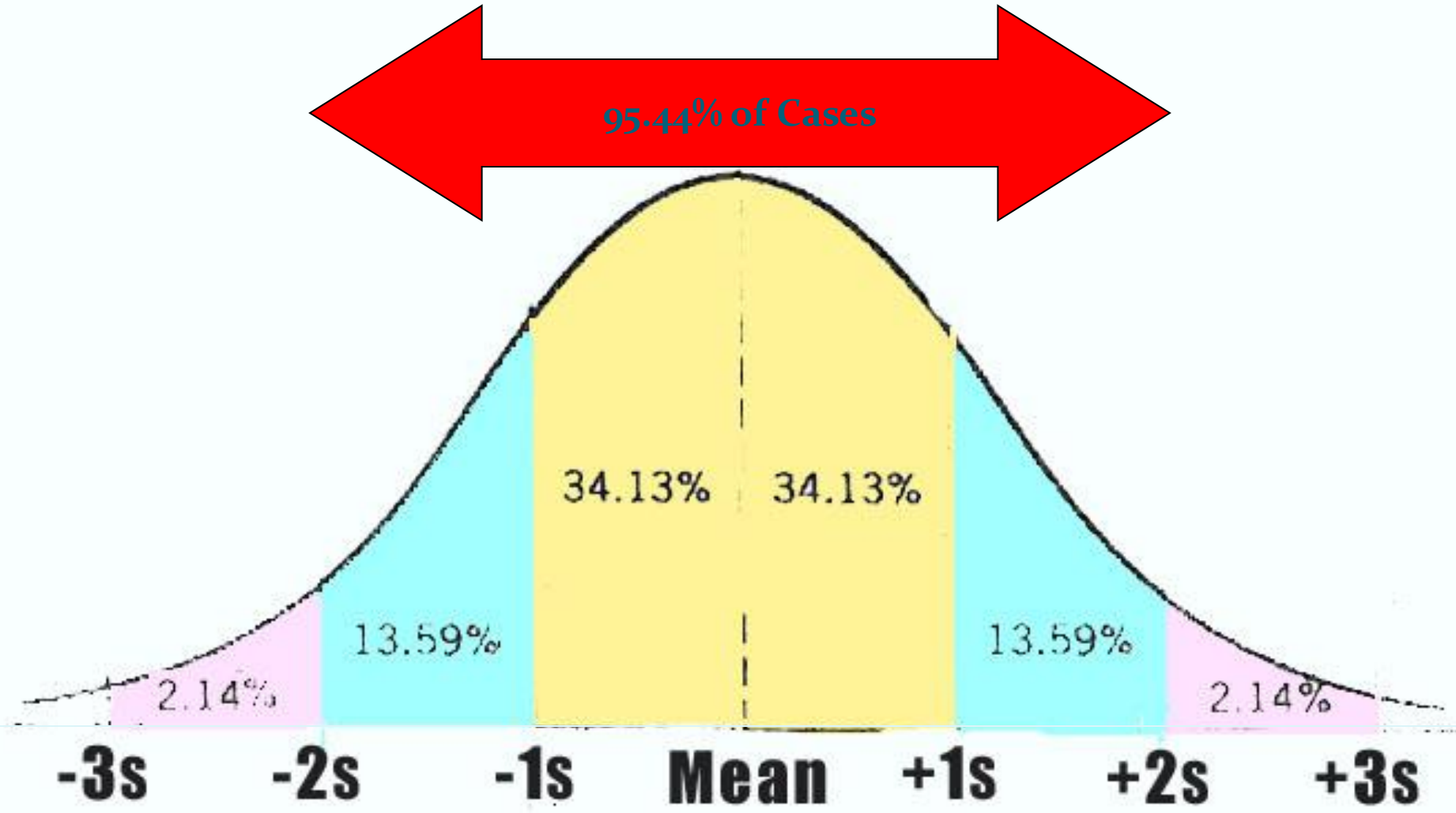
Percent of Values Within One Standard Deviations



NORMAL DISTRIBUTION

- (b)The area between two standard deviations on either side of the mean ($\bar{x} \pm 2\sigma$) will cover most of the values, i.e., approximately 95 per cent of the values.

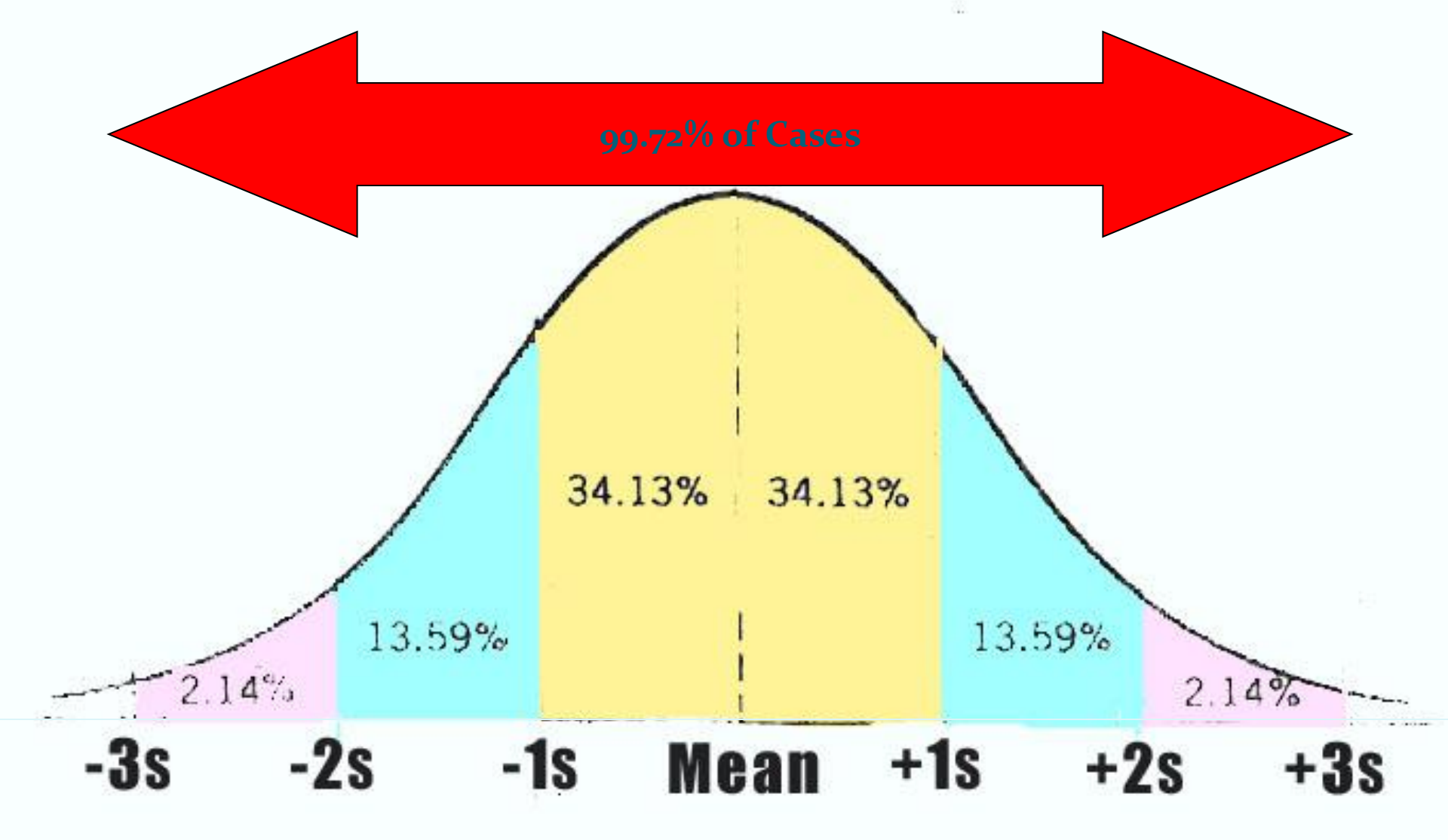
Percent of Values Within Two Standard Deviations



NORMAL DISTRIBUTION

- (c) The area between $(\bar{x} \pm 3\sigma)$ will include 99. 7 per cent of the values.

Percent of Values Within Three Standard Deviations



- These limits on either side of the mean are called "confidence limits“
- Supposing we are considering the 95 per cent confidence limits ($\bar{x} \pm 2\sigma$). When we say this, we mean that 95 per cent of the area of the normal curve, and hence 95 per cent of the values in the distribution will be included between the limits $\bar{x} \pm 2\sigma$. Therefore, the probability of a reading falling outside the 95 per cent confidence limits is 1 in 20 (P = 0.05).

Question

- Blood pressure of 150 doctors was recorded. The mean BP was found to be 127 mmHg (systolic). The standard deviation was calculated to be 6 mmHg.
- If blood pressure is normally distributed, then how many doctors will have systolic blood pressure above 133 mmHg?

Standard normal curve

- Although there is an infinite number of normal curves depending upon the mean and standard deviation, there is only one standardized normal curve, which has been devised by statisticians to estimate easily the area under the normal curve, between any two ordinates.

Standard normal curve

- The standard normal curve is a

- smooth,
- bell-shaped,
- perfectly symmetrical curve, based on an infinitely large number of observations.

Standard normal curve

- The total area of the curve is 1;
- Its mean is zero; and
- Its standard deviation is 1.
- The mean, median and mode all coincide.

Standard normal curve

- The distance of a value (x) from the mean ($\bar{X}$) of the curve in units of standard deviation is called "relative deviate or standard normal variate" and is usually denoted by Z.

Standard normal curve

- The standard normal deviate or Z is given by the formula

$$Z = \frac{x - \bar{x}}{\sigma}$$

Standard normal curve

- A random variable (x) is said to have been standardized when it has been adjusted so that its mean is zero and its standard deviation is 1.

Standard normal curve

- Standardization can be effected by subtracting the mean of x, from $\bar{x}$ and dividing the resulting difference by σ, the standard deviation of x.
- Thus $\frac{x - \bar{x}}{\sigma}$ is a standardized variable.

Standard normal curve

- An important concept of mathematical statistics is that the new variate "Z" like the variate "x" also follows a normal distribution. The mean of the transformed distribution is zero (0), and the standard deviation (σ) is 1.

Standard normal curve

- Areas under the standard normal curve are frequently needed. They have been computed for values of different relative deviate $\left(\frac{x - \bar{x}}{\sigma} \right)$

Standard normal curve

Relative deviate (z) (x.-x)	Proportion of area from middle of the curve of designated deviation
0.00	0.0000
0.50	.1915
1.00	.3413
1.50	.4332
2.00	.4772
3.00	.4987
4.00	.49997

(Example)

- Let us suppose, the pulse of a group of normal healthy males was 72, with a standard deviation of 2.
- What is the probability that a male chosen at random would be found to have a pulse of 80 or more ?

Relative deviate

$$\text{The relative deviate } (z) = \frac{x - \bar{x}}{\sigma}$$

$$= \frac{80 - 72}{2}$$

$$= 8/2 = 4$$

Relative deviate

- The area of the normal curve corresponding to a deviate 4=0.49997 . Since we are dealing with only half the total area (i.e., 0.5) the area beyond 0.49997 is equal to 0.5 - 0.49997 or 0.00003.
- Therefore, the probability is that only 3 out of 100,000 individuals would likely have a pulse rate of 80 or higher.

MEASURES OF SHAPE OF DISTRIBUTION

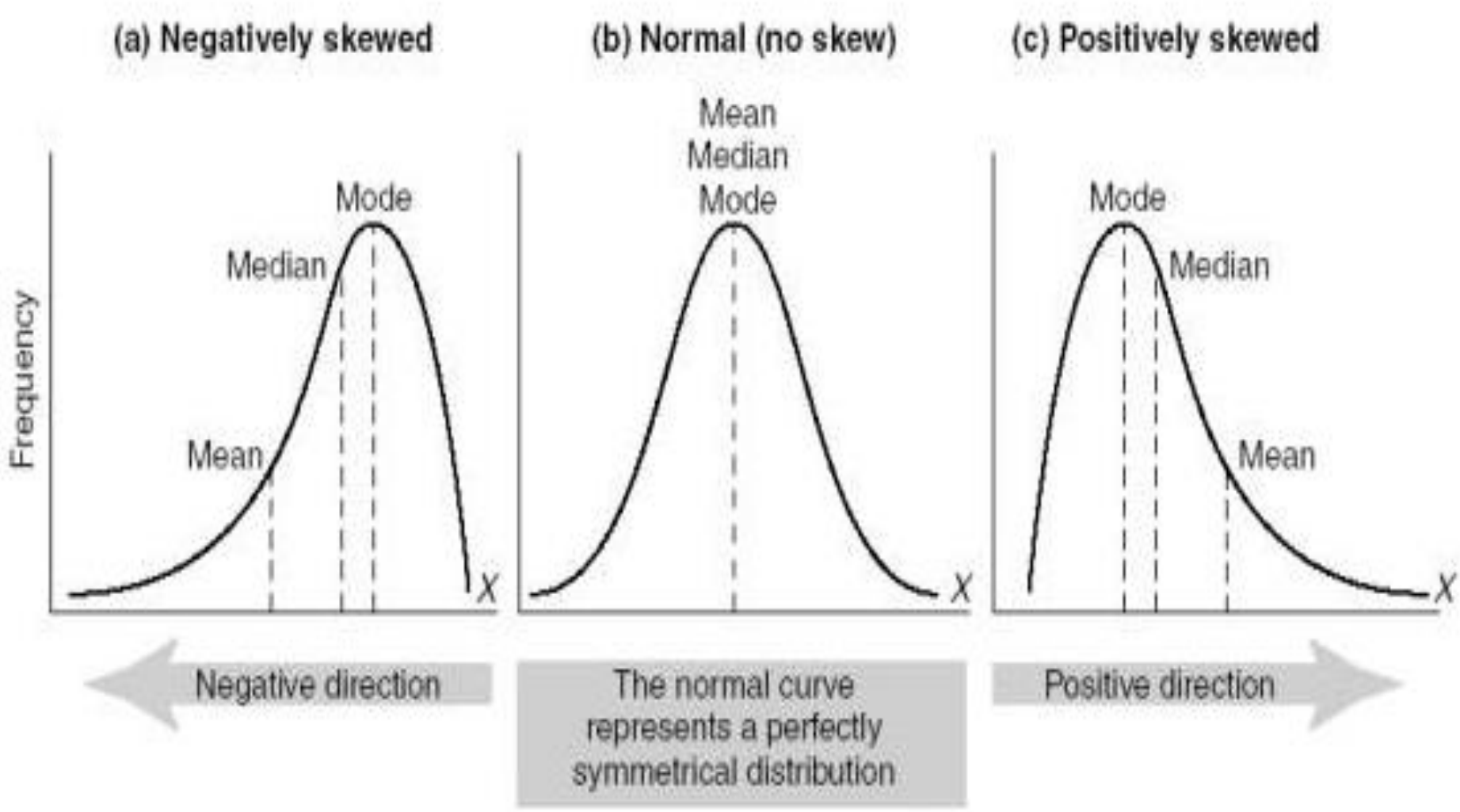
Skewness, Kurtosis

- Skewness (S_k), Pearsonian coefficient, is a measure of asymmetry of a distribution around its mean.
- Kurtosis characterizes the relative peakedness or flatness of a distribution compared with the normal distribution.

SKEWED DISTRIBUTION

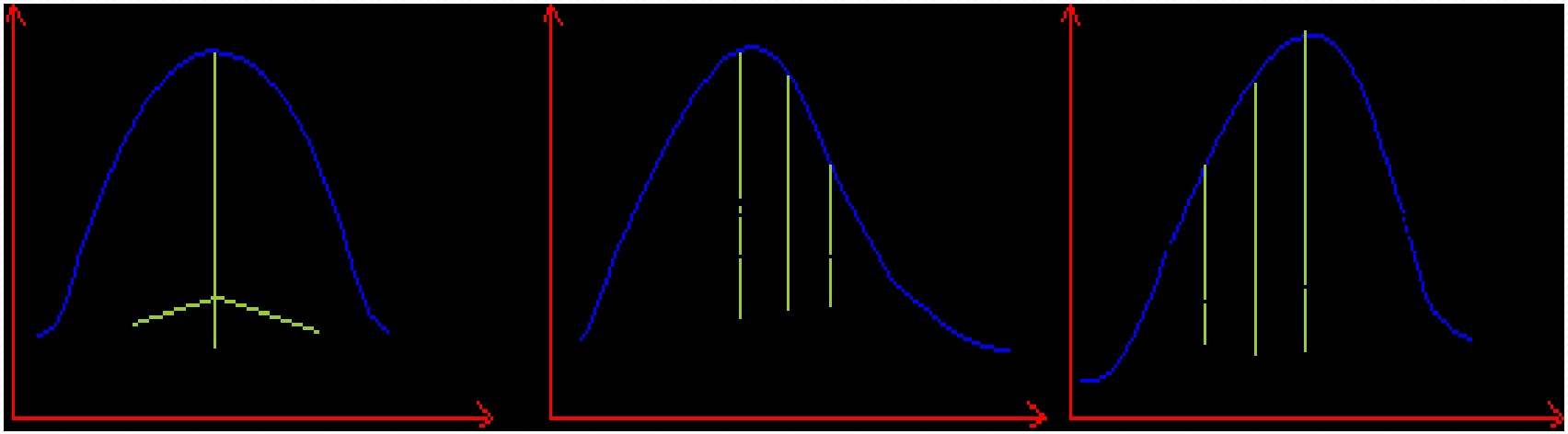
- When frequency distribution or frequency curve is not symmetrical about the peak, it is said to be "skewed"(asymmetrical)
- In other words one tail of the curve will be longer than the other. This skewness can be either to the right or to the left of the peak

Skewed distribution Curves

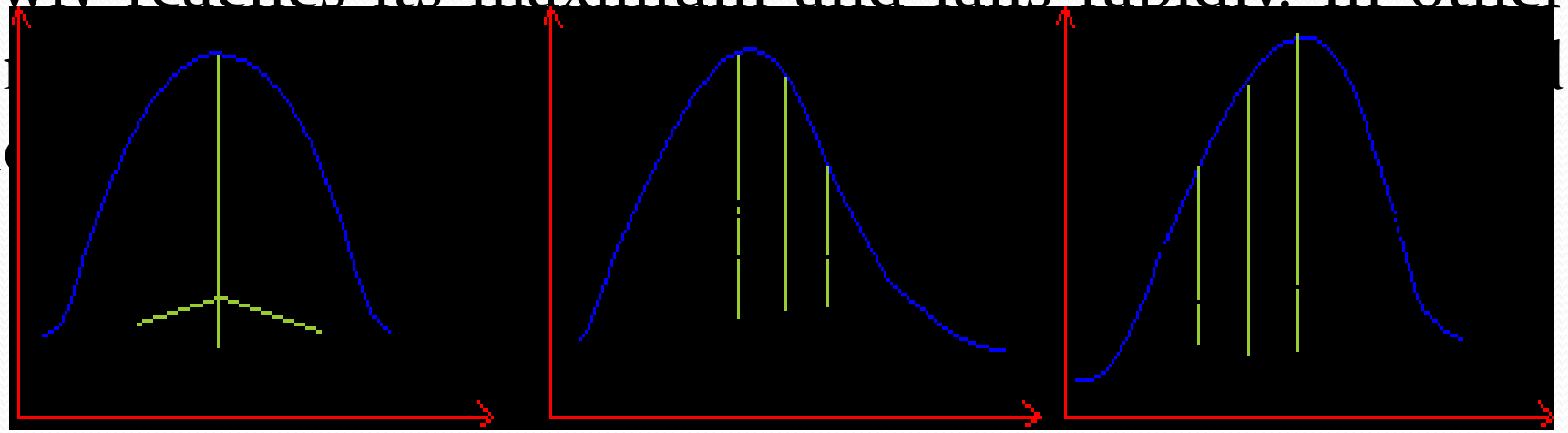


Symmetrical and Asymetrical distribution

- ✓ For a symmetrical distribution, the values, of equal distances on either side of the mode, have equal frequencies. Thus, the mode, median and mean - all coincide. Its curve rises slowly, reaches a maximum (peak) and falls equally slowly (Fig. 1).

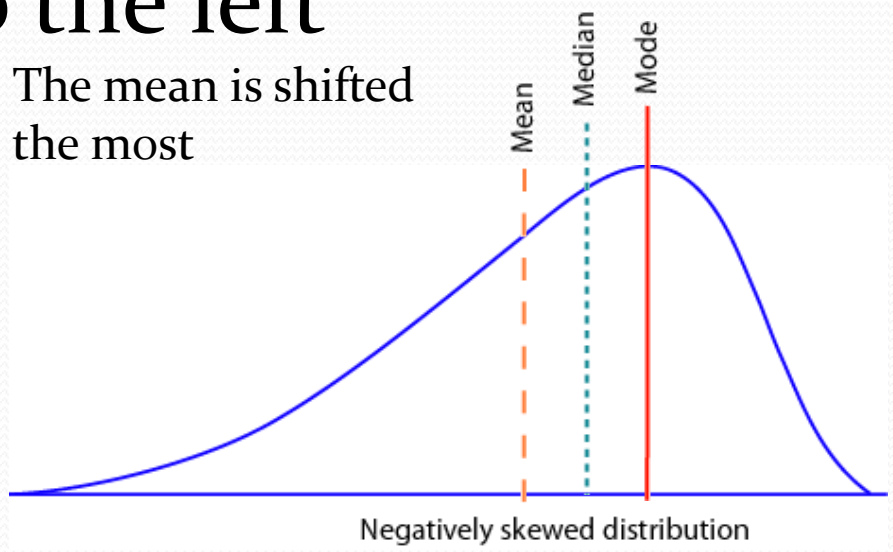


- ✓ But for a skewed distribution, the mean, mode and median do not coincide. Skewness is positive or negative as per the positions of the mean and median on the right or the left of the mode.
- ✓ A positively skewed distribution (Fig.2) curve rises rapidly, reaches the maximum and falls slowly. In other words, the tail as well as median on the right-hand side.
- ✓ A negatively skewed distribution curve (Fig.3) rises slowly reaches its maximum and falls rapidly. In other wo
sid



Skewed distributions

- Skew is always toward the direction of the longer tail
- Positive if skewed to the right
- Negative if to the left



Skewed distributions

- Because the mean is shifted so much, it is not the best estimate of the average score for skewed distributions
- The median is a better estimate of the center of skewed distributions
  - It will be the central point of any distribution
  - 50% of the values are above and 50% below the median

QUESTION

- The weights of 4^{th} year MBBS class were recorded. The analysis showed
- mean=50Kg, mode=54Kg, median=52kg.
- This distribution will be labelled as:?

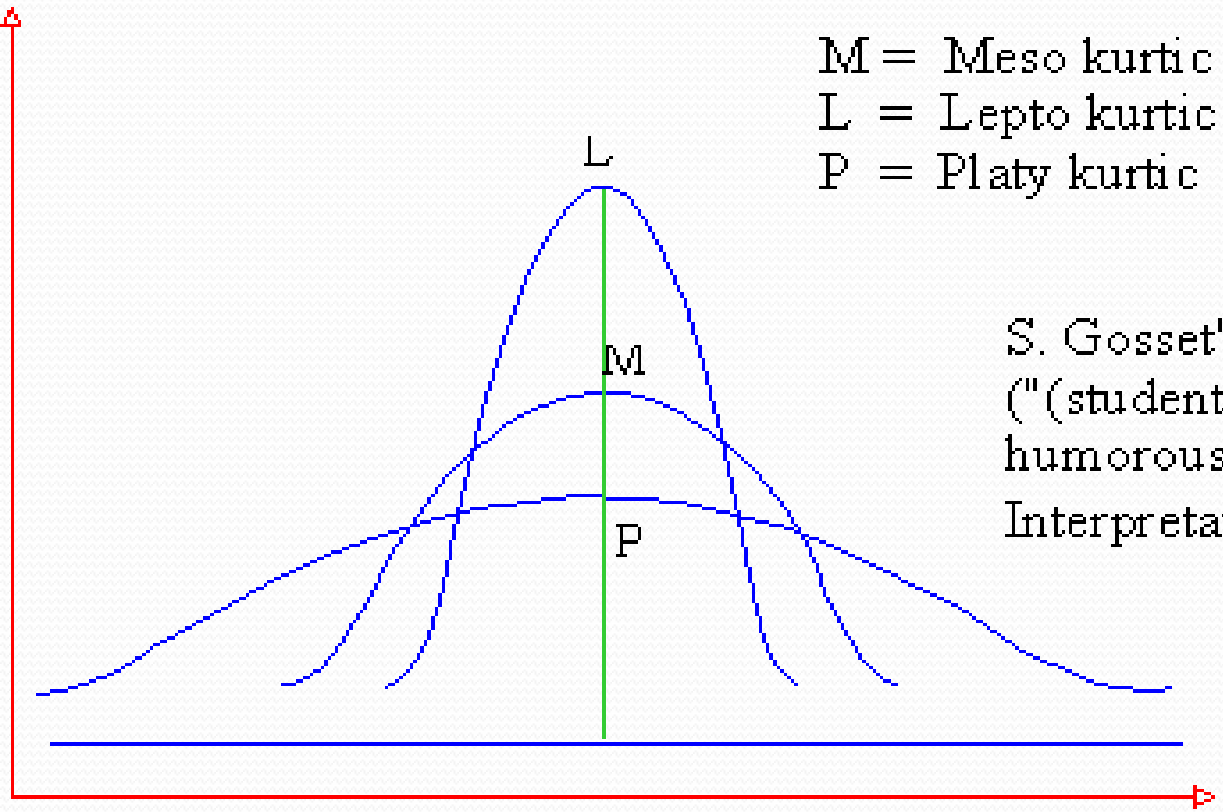
QUESTION

- In a study random blood sugar level of 150 medical students was check. Various values were as
- mean= 130 mg/100L, median 125, mode 120.
- This distribution is called as:?

KURTOSIS

- It is a term to show the degree of peakedness of a distribution.
  1. If the peak of the distribution is relatively high, the distribution is "LEPTO-KURTIC".
  2. On the other hand the curve is flat toped the distribution is "PLATY – KURTIC".
  3. The normal curve is neither very peaked nor very flat topped. This curve is called "MESO-KURTIC". It is also known as normal curve

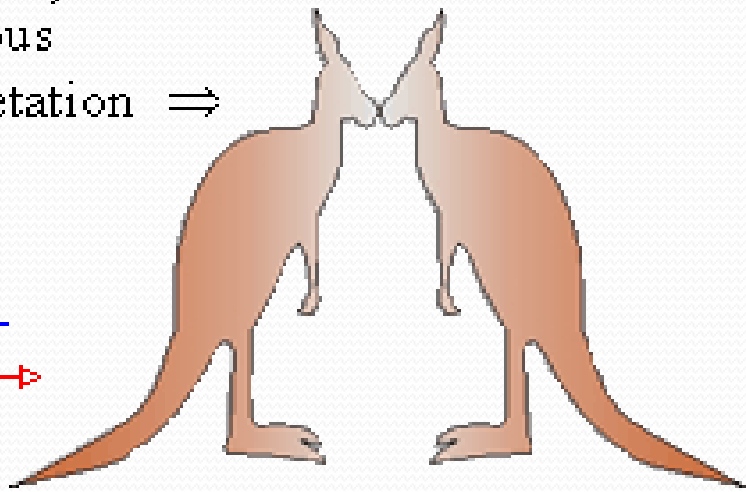
Kurtosis



M = Meso kurtic
L = Lepto kurtic
P = Platy kurtic

(Platy kurtic)

S. Gosset's ("(student)" humorous Interpretation ⇒



Long tailed Kangaroos

(Lepto kurtic)


Thank You