

Normal distribution **(The normal distribution curve)**

By
Dr. Asaad Al-Yassen
Assist. Professor of dermato-epidemiology
& occupational health
Department of Community Medicine
University of Basrah

Learning objectives:

At the end of the lecture they should be able to:

1. Define the normal distribution curve and skewed distribution.
2. Know the properties of normal distribution curve.
3. Know the advantages or the applications of normal distribution curve.
4. Define the normal range and Z-Score.

The normal distribution curve

It is one form of symmetrical distribution. It has very great an important in statistical theory, because the distribution of many medical measurements in the population approximate the normal distribution, and many measurable characteristics (continuous variables) occurring in the nature show this kind of distribution, e.g. blood pressure, hemoglobin level, height, weight, body temperature, and serum uric acid.....

So that many continuous variables have distribution that are bell shaped called **normally distributed variables**.

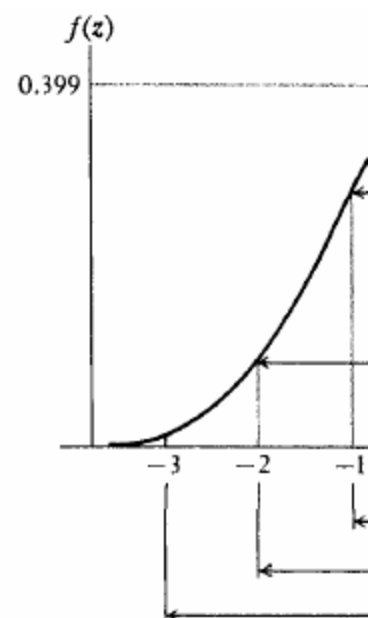
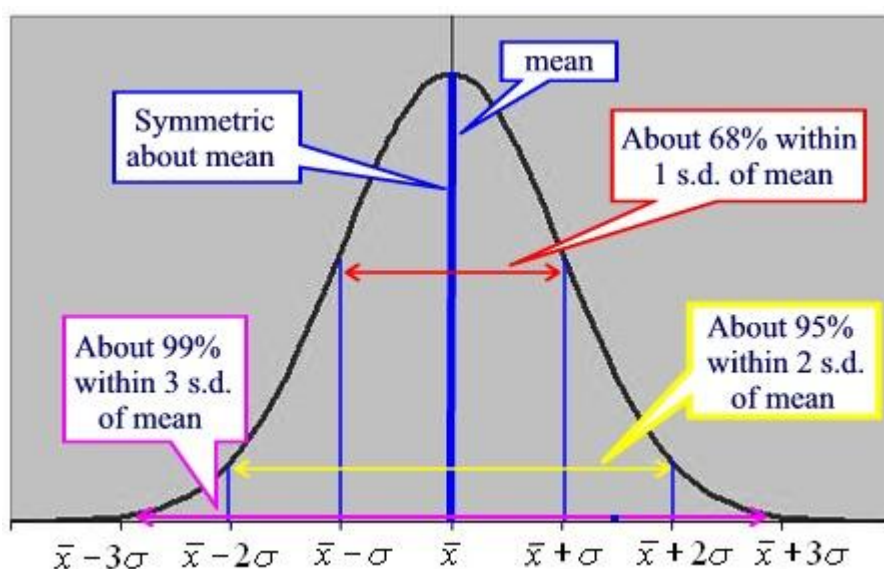
Normal distribution is also known as **the bell curve** or **the Gaussian distribution**.

The shape and position of the normal distribution curve depends on two parameters, **the mean** and **the standard deviation**.

The spread of a normal distribution is controlled by **standard deviation**.

Characteristics of normal distribution:

1. The normal distribution curve is bell shape.
2. The mean, median, and mode are equal and located at the center of the distribution.
3. The normal distribution curve is unimodal (single mode)
4. The curve is symmetrical about the mean, 50% of the values lie above the mean (to the right), and 50% of the values lie below the mean (to the left).
5. The curve is continuous.
6. The curve never touches the X axis.
7. The total area under the normal curve that lies within:
 - * one standard deviation of the mean is approximately 68%.
 - * two standard deviations of the mean is approximately 95%.
 - * three standard deviations of the mean is approximately 99.7%.



Advantages of normal distribution:

1. We can theoretically calculate how many observations will lie within a given distance of the mean in term of multiples of standard deviation.(not in term of actual unit)

$$\bar{X} \pm K Sd$$

*proportion of observations that lie between $\bar{X} \pm 1 Sd$ are equal 68%.

*proportion of observations that lie between $\bar{X} \pm 2 Sd$ are equal 95%.

i.e. 95% of the observations lie between $\bar{X} \pm 2 Sd$

This can be considered as **a normal range** or **normal status** for a measurement that follows a normal distribution.

i.e. only 5% of the observations will differ from the mean by more than twice the standard deviation (outside the normal range).

*proportion of observations that lie between $\bar{X} \pm 3 Sd$ are equal 99.7%.

Accordingly, we can express, how far an observation differ from the mean in term of multiple standard deviations by what we called **Z-Score**.

Z-Score: The location of any observation can be expressed in terms of how many standard deviations lies above or below the mean of the distribution.

$$\text{Z-Score} = \frac{X - \bar{X}}{Sd}$$

If it lies above the mean $\rightarrow$ +ve Z-Score.

If it lies below the mean $\rightarrow$ -ve Z-Score.

2. To determine the normal range of biological and medical variables whose distribution is known to be a proposed normal.
There are disease entities that are based almost exclusively on a single parameter.

e.g. iron deficiency anaemia → based on hemoglobin level
hypertension → blood pressure
diabetes mellitus → glucose level

So that the intervention or therapy is aim to bring the level back to normal range, and this will help to reduce the risk of disease.

Skewed distribution:

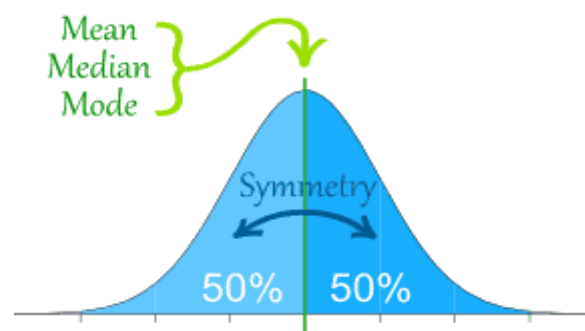
- The normal distribution is a continuous, bell shaped distribution of a variable.
- If the data values are evenly distributed about the mean, the distribution is said to be symmetrical.
- If the majority of the data values fall to the **left**, or to the **right** of the mean, the distribution is said to be **skewed**.

The Normal Distribution has No Skew

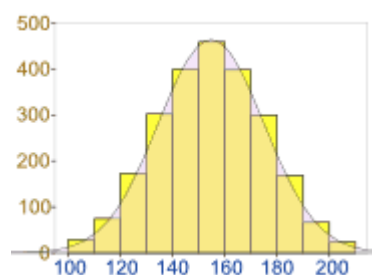
A [Normal Distribution](#) is not skewed.

It is perfectly symmetrical.

And the Mean is exactly at the peak.



Mean = Median= Mode

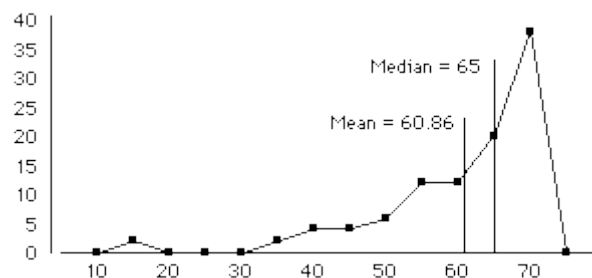
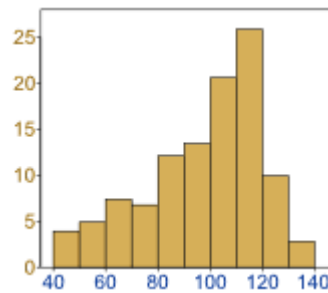


Left skewed distributions:

When the majority of the data values fall to the right of the mean, the distribution is said to be **negatively** or **left skewed**.

The mean is to the left of the median and the mean and the median are to the left of the mode.

$$\text{Mean} < \text{Median} < \text{Mode}$$



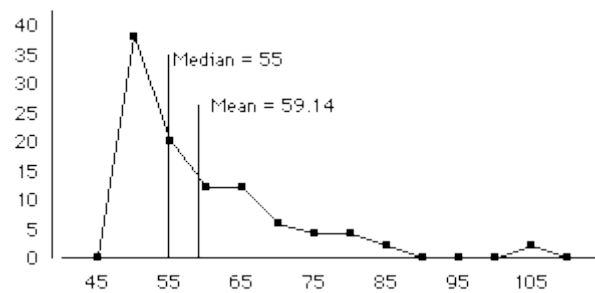
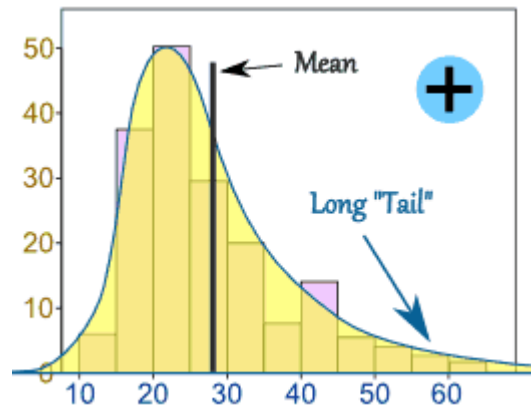
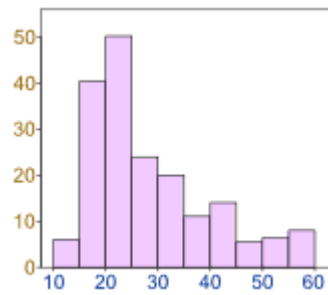
This distribution has a negative skew. The median is larger than the mean.

Right skewed distributions:

When the majority of the data values fall to the left of the mean, the distribution is said to be **positively** or **right skewed**.

The mean is to the right of the median and the mean and the median are to the right of the mode.

$$\text{Mean} > \text{Median} > \text{Mode}$$



This distribution has a positive skew. Note that the mean is larger than the median.

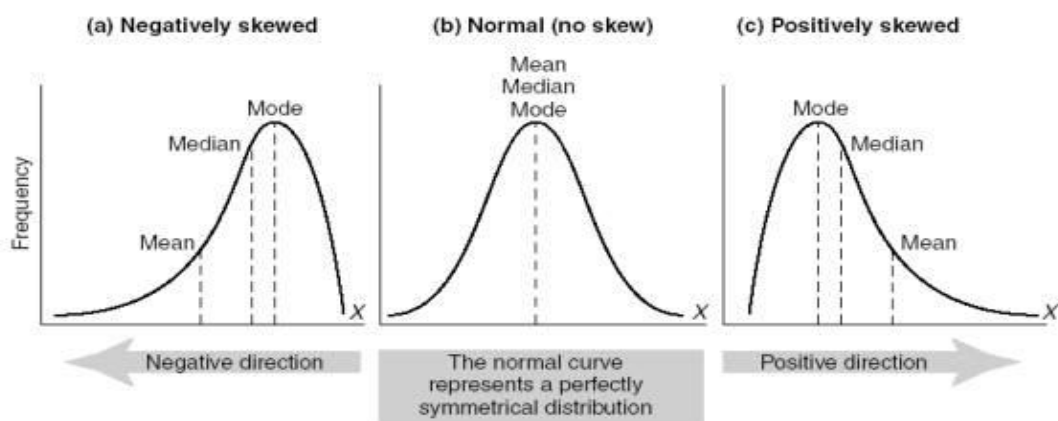


FIGURE 15.6 Examples of normal and skewed distributions