

SNS COLLEGE OF TECHNOLOGY

**An Autonomous Institution
Coimbatore-35**



DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

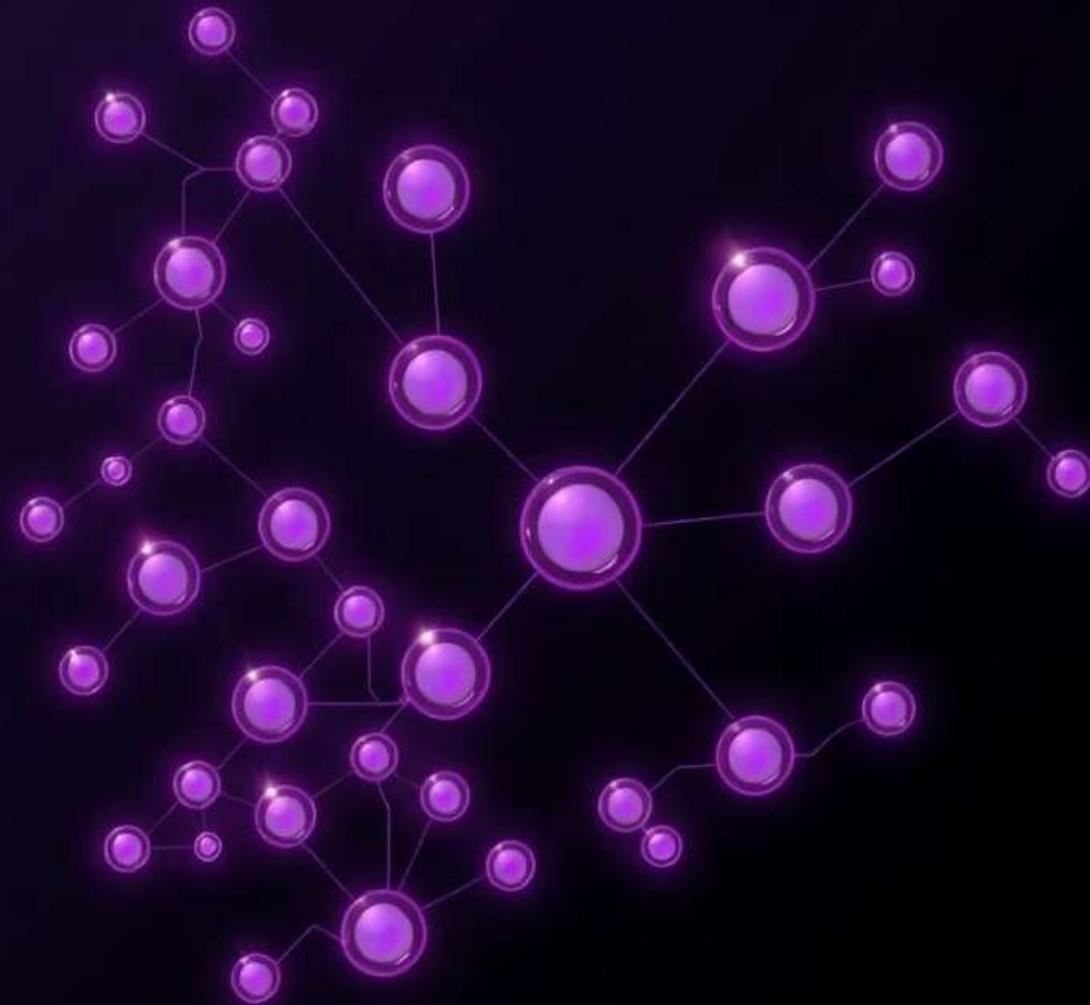
23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT II – NORMAL DISTRIBUTIONS

EMPATHY:

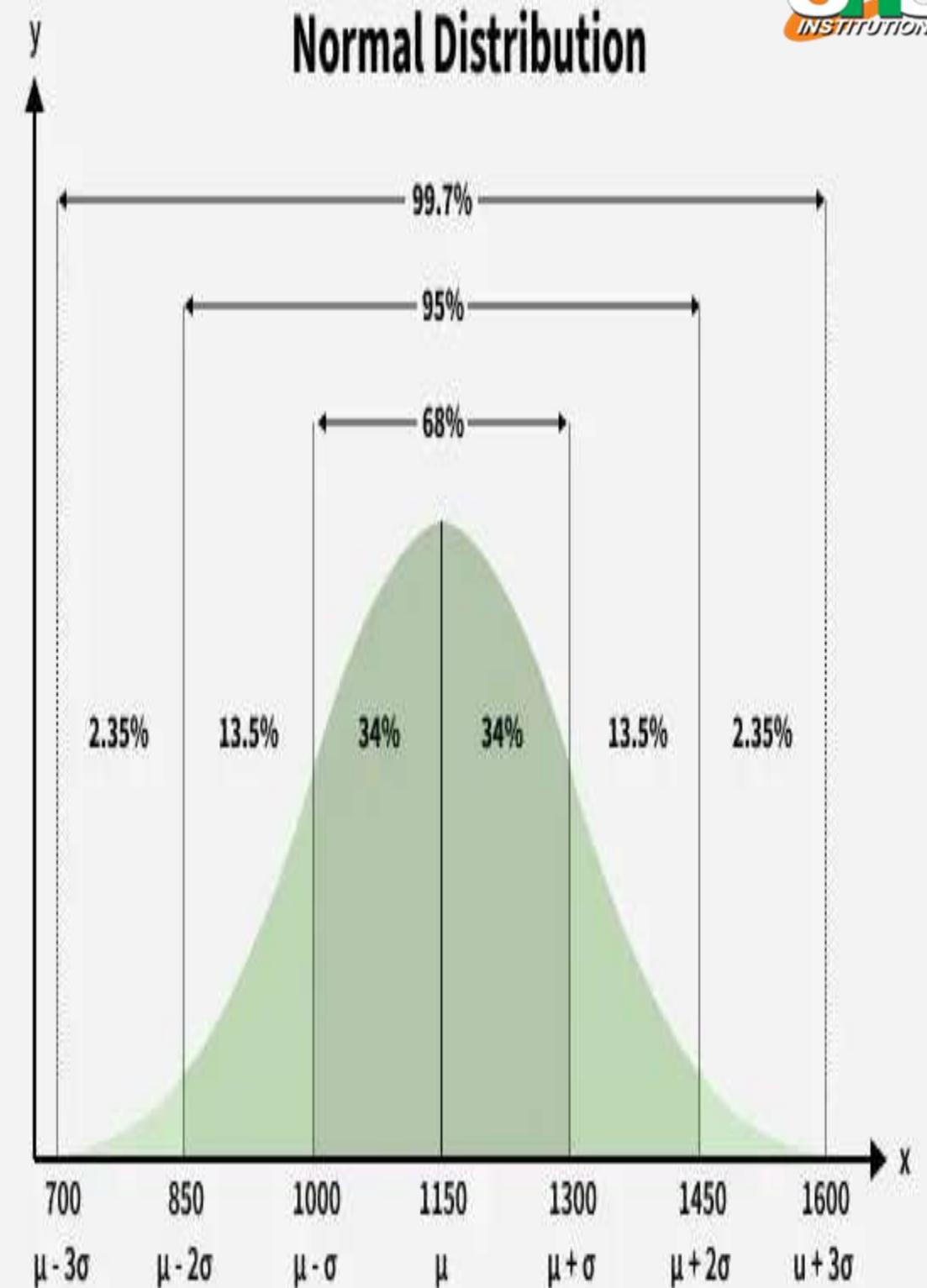
- ❖ Normal Distribution is the most common or normal form of distribution of Random Variables, hence the name "normal distribution."
- ❖ It is also called the Gaussian Distribution in Statistics or Probability.
- ❖ We use this distribution to represent a large number of random variables.
- ❖ It serves as a foundation for statistics and probability theory.



Normal distribution

DEFINE:

- ❖ It is a continuous probability distribution that is symmetric about the mean, depicting that data near the mean are more frequent in occurrence than data far from the mean.
- ❖ As shown in Fig 1, the distribution is symmetric about its center, which is the mean (0 in this case).
- ❖ This symmetry means that events equidistant from the mean have equal probabilities.
- ❖ The density is highest near the mean, resulting in lower probabilities for values farther away from it.



- ❖ We define Normal Distribution as the probability density function of any continuous random variable for any given system.
- ❖ Now for defining Normal Distribution suppose we take $f(x)$ as the probability density function for any random variable X .
- ❖ The area under the curve of a normal distribution is always 1.
- ❖ The curve traced by the upper values of the Normal Distribution is in the shape of a Bell, hence Normal Distribution is also called the "Bell Curve".

Normal Distribution Formula

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where,

x is Random Variable
 μ is Mean
 σ is Standard Deviation

The formula for the probability density function of the Normal Distribution (Gaussian Distribution) is added below:

IDEATE:

Normal Distribution Characteristics

Symmetry: The normal distribution is symmetric around its mean. This means the left side of the distribution mirrors the right side.

Mean, Median, and Mode: In a normal distribution, the mean, median, and mode are all equal and located at the center of the distribution.

Bell-shaped Curve: The curve is bell-shaped, indicating that most of the observations cluster around the central peak and the probabilities for values further away from the mean taper off equally in both directions.

Standard Deviation: The spread of the distribution is determined by the standard deviation. About 68% of the data falls within one standard deviation of the mean, 95% within two standard deviations, and 99.7% within three standard deviations.

PROTOTYPING:

Normal Distribution Examples

We can draw a Normal Distribution for various types of data that include,

- ❖ Distribution of Height of People.
- ❖ Distribution of Errors in any Measurement.
- ❖ Distribution of Blood Pressure of any Patient, etc.

Normal Distribution Curve

In a Normal Distribution, a random variable (X) is a numerical outcome of a process that follows this distribution.

- ❖ The values of X are not fixed but instead vary according to the distribution's properties, where:
- ❖ The variable is continuous (can take any real value within a range).
- ❖ The distribution is defined by its mean (μ) - the peak of the curve and standard deviation (σ) - which controls the spread of the curve.
- ❖ The Normal Distribution Curve (also called the Bell Curve or Gaussian Curve) is the graphical representation of this distribution, showing:
 - ❖ Symmetry around the mean (μ).
 - ❖ 68-95-99.7% Rule (Empirical Rule) for data spread.
 - ❖ Asymptotic tails (the curve never touches the x-axis but extends infinitely).