



# SNS COLLEGE OF ENGINEERING



Kurumbapalayam(Po), Coimbatore – 641 107

Accredited by NAAC-UGC with 'A' Grade

Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

## Department of Information Technology

Course Name – 23ADT202 Fundamental of Data  
science and Analytics

II Year / IV Semester

Unit 2 – Descriptive Analytics

Normal Distributions





# What is a Normal Distribution?

A normal distribution is a probability distribution that is symmetric about its mean. It is bell-shaped, with most data points clustered around the mean. It appears in many real-world phenomena like heights, test scores, and measurement errors.

This distribution is fundamental to statistics and data science.



# Key Characteristics of Normal Distributions



1. Symmetric around the mean.
2. Mean, median, and mode are equal.
3. Bell-shaped curve, with data clustered around the mean.
4. The Empirical Rule: 68-95-99.7% for 1, 2, and 3 standard deviations.



# Probability Density Function (PDF)



The normal distribution is defined by a probability density function (PDF). It is characterized by two parameters: mean ( $\mu$ ) and variance ( $\sigma^2$ ). The formula for the PDF describes the likelihood of a value occurring. It is crucial for calculating probabilities in statistics.



# Probability Density Function (PDF)



The normal distribution is defined by a probability density function (PDF). It is characterized by two parameters: mean ( $\mu$ ) and variance ( $\sigma^2$ ). The formula for the PDF describes the likelihood of a value occurring. It is crucial for calculating probabilities in statistics.



# Standard Normal Distribution



The standard normal distribution is a special case of the normal distribution. It has a mean ( $\mu$ ) of 0 and a standard deviation ( $\sigma$ ) of 1. Any normal distribution can be transformed into the standard normal distribution. This is done using the Z-score formula.



# Visualizing a Normal Distribution



The bell curve is used to represent a normal distribution graphically. Data is concentrated around the mean, with tails approaching zero. The area under the curve represents total probability, which sums to 1. This visualization helps understand the spread of data.



# Applications of Normal Distributions



In education, normal distributions model test scores for grading curves.  
In healthcare, biological measurements like blood pressure follow normal distributions.  
In finance, investment returns often assume normality.  
Quality control uses it to assess manufacturing deviations.





# Central Limit Theorem



The Central Limit Theorem states that the sampling distribution of the sample mean approaches normality.

This holds true regardless of the original data distribution.

As the sample size increases, the distribution of the mean becomes more normal.

This is a powerful concept in statistical inference.



# Outlier Detection in Normal Distributions



Data points beyond 3 standard deviations ( $\sigma$ ) from the mean are considered outliers. These outliers are rare and indicate significant deviations from the norm. They can distort the analysis, so identifying them is important. A normal distribution is useful for detecting unusual values.



# Conclusion and Exercise



Normal distributions are essential in understanding real-world data. Key concepts include symmetry, bell-shape, and the empirical rule. Applications range from education to quality control. Exercise: Plot a normal distribution with  $\mu = 50$ ,  $\sigma = 10$ ; calculate Z-score for  $X = 60$ .



**THANK YOU**