ASSIGNMENT (QUESTION -2)

Take one Domain and draw the graph (Normal distribution) (Empirical rule)

Introduction

In statistics and data analysis, the normal distribution is one of the most important probability distributions used to represent real-world data. It describes how data values are distributed around a mean or average value. The normal distribution is also known as the **bell-shaped curve** because of its symmetrical shape, where most values are concentrated around the center and fewer values appear at the extremes.

The empirical rule, also called the **68–95–99.7 rule**, explains how data is distributed in a normal distribution. It describes the percentage of data values that fall within one, two, and three standard deviations from the mean. The normal distribution is widely used in various domains such as education, healthcare, finance, and business.

In this assignment, the domain selected is **student test scores**, and the normal distribution graph is used to analyze how students' marks are distributed around the average score using the empirical rule.

Domain Selection: Student Test Scores

The education domain provides a good example of normal distribution because student test scores often follow a bell-shaped pattern. Most students score around the average marks, while fewer students achieve very high or very low scores.

Why Test Scores Follow Normal Distribution

- Students have varying levels of understanding and preparation.

- Most students perform at an average level.

- Only a few students score extremely high or extremely low.

- External factors such as study time, difficulty level, and learning ability influence performance.

Because of these characteristics, test scores typically form a symmetric distribution around the mean.

Understanding Normal Distribution

A normal distribution is defined by two important parameters:

1. Mean (μ)

The mean represents the average value of the dataset. In test scores, it represents the average marks obtained by students.

2. Standard Deviation (σ)

Standard deviation measures how far data values spread from the mean. A small standard deviation means scores are close to the average, while a large standard deviation indicates wide variation.

Characteristics of Normal Distribution

- Bell-shaped and symmetric curve.
- Mean, median, and mode are equal.
- Most values cluster around the mean.
- Probability decreases as values move away from the mean.

Empirical Rule (68–95–99.7 Rule)

The empirical rule explains the percentage of data within standard deviations from the mean.

1. 68% Rule (One Standard Deviation)

Approximately **68% of data** falls within one standard deviation from the mean ($\mu \pm 1\sigma$).

Example in Test Scores

If the average score is 70 and standard deviation is 10:

- Most students score between 60 and 80.
- These students represent average performers.

This region forms the central part of the bell curve.

2. 95% Rule (Two Standard Deviations)

About **95% of data** falls within two standard deviations from the mean ($\mu \pm 2\sigma$).

Example

- Scores range between 50 and 90.
- Includes average and above/below average students.

This covers almost all students except extreme cases.

3. 99.7% Rule (Three Standard Deviations)

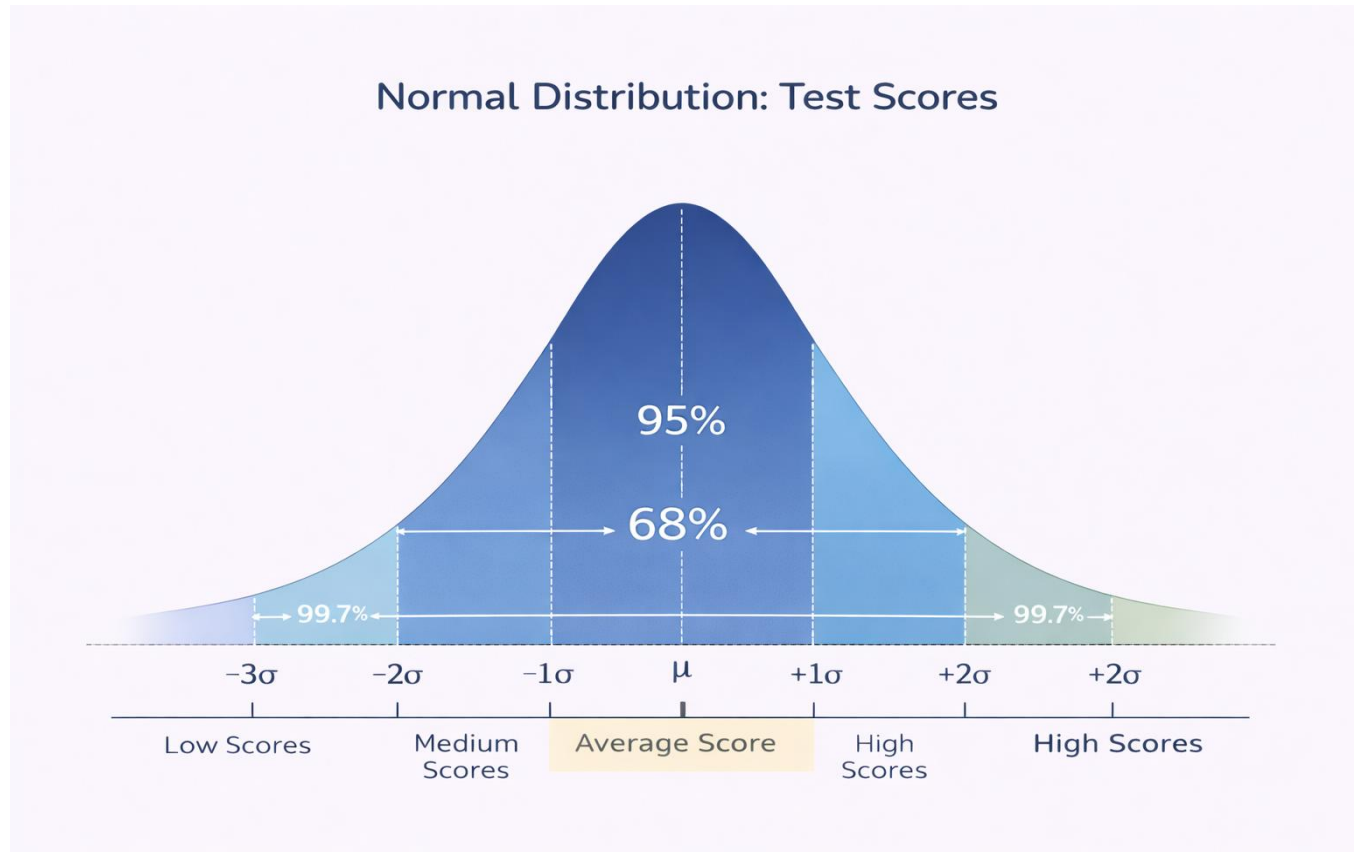Nearly **99.7% of data** falls within three standard deviations from the mean ($\mu \pm 3\sigma$).

Example

- Scores range between 40 and 100.

- Includes nearly all students.
- Only very few students fall outside this range.

This represents the entire distribution.

Graph Representation of Normal Distribution

The normal distribution graph for test scores is a bell-shaped curve centered at the mean.

Components of the Graph

X-axis

Represents student scores ranging from low to high.

Y-axis

Represents frequency or number of students.

Mean (μ)

Located at the center of the curve representing average score.

Standard Deviations

Marked as:

- $\mu - 1\sigma$, $\mu + 1\sigma$

- $\mu - 2\sigma$, $\mu + 2\sigma$

- $\mu - 3\sigma$, $\mu + 3\sigma$

The curve is symmetric around the mean.

Interpretation of Student Performance Using Empirical Rule

Using the normal distribution graph, student performance can be classified as follows:

Low Scores (Below -2σ)

- Few students

- Poor performance

- Require additional support

Medium Scores (Between -1σ and $+1\sigma$)

- Majority of students

- Average performance

High Scores (Above $+1\sigma$)

- Fewer students

- Above average performance

Extreme Scores (Beyond $\pm3\sigma$)

- Very rare cases

This classification helps educators understand performance distribution.

Applications of Normal Distribution in Education

Normal distribution is widely used in the education sector.

1. Performance Analysis

Helps analyze student performance levels.

2. Grading Systems

Used for grading on a curve.

3. Exam Difficulty Evaluation

Helps determine if exam difficulty is balanced.

4. Predicting Student Outcomes

Helps predict future performance trends.

5. Identifying Outliers

Helps identify exceptionally high or low performers.

Advantages of Using Normal Distribution

- Simple and easy interpretation.

- Useful for prediction and decision-making.

- Represents real-world data effectively.

- Helps identify patterns in large datasets.

Limitations of Normal Distribution

- Not all data follows normal distribution.

- Sensitive to extreme values.

- Requires large datasets for accuracy.

Despite limitations, it remains widely used in statistical analysis.

Conclusion

The normal distribution provides an effective way to represent and analyze data in real-world scenarios. In the domain of student test scores, the bell-shaped curve demonstrates how most students score around the average, while fewer students achieve extreme scores. The empirical rule explains the percentage distribution of data within one, two, and three standard deviations from the mean, helping to understand student performance patterns.

The use of normal distribution in education supports grading, performance analysis, and decision-making. Understanding the empirical rule allows educators and researchers to interpret data accurately and identify trends. Therefore, normal distribution and the empirical rule play an important role in statistical analysis and real-world data interpretation.