Assignment 2:

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

Name : Akshay Kumar V Padmashali

USN : 01SU24CS015

Branch : CSE

Section : A

1. Introduction

Statistics plays a vital role in analyzing real-world data across multiple domains such as education, healthcare, finance, sports, and engineering. One of the most important and widely used probability distributions in statistics is the **Normal Distribution**, also known as the Gaussian distribution. It helps describe how data values are spread around the mean in many natural and social phenomena.

In this report, we explore the concept of the Normal Distribution and the Empirical Rule using the domain of **Student Exam Scores**. Exam scores often follow a bell-shaped distribution where most students score around the average, fewer score extremely high or extremely low, and the spread follows a predictable pattern. This makes it an ideal example for understanding statistical concepts such as mean, standard deviation, and probability distribution.

2. Understanding Normal Distribution

2.1 Definition

A Normal Distribution is a continuous probability distribution that is symmetric about the mean. It is characterized by a bell-shaped curve where:

- The highest point represents the mean.
- The curve decreases equally on both sides.
- The total area under the curve equals 1.

2.2 Key Characteristics

1. Symmetrical about the mean
2. Mean = Median = Mode
3. Defined by two parameters:
    a. Mean (μ)
    b. Standard Deviation (σ)

2.3 Mathematical Formula

The probability density function (PDF) of a normal distribution is:

$f(x)=1σ2πe−(x−μ)22σ2f(x)$ = $\frac{1}{\sigma \sqrt{2\pi}}$ $e^{-\frac{(x-\mu)^2}{2\sigma^2}}f(x)=σ2π$ 1 $e−2σ2(x−μ)2$

Where:

- μ = Mean
- σ = Standard Deviation
- π ≈ 3.1416
- e ≈ 2.718

3. Domain Selected: Student Exam Scores

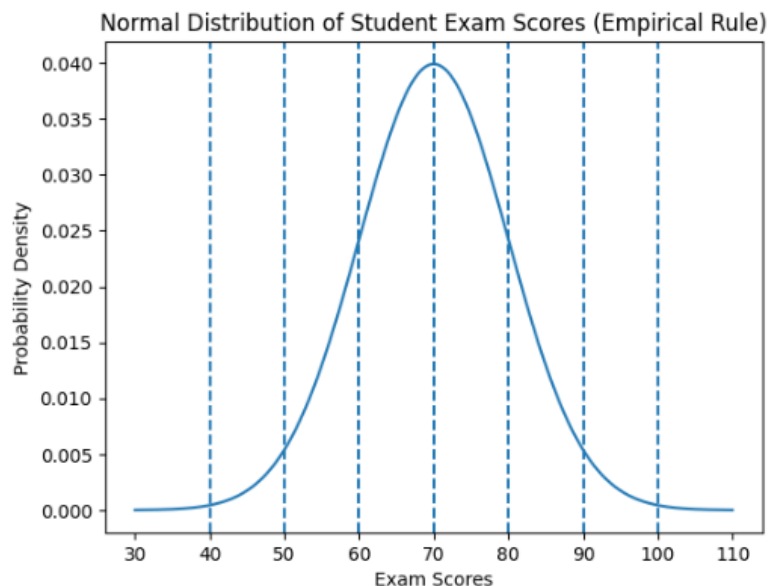In this report, we consider a class where:

- **Mean score (μ) = 70**
- **Standard Deviation (σ) = 10**

This means:

- The average student scored 70.
- Most students' scores deviate about 10 marks from the average.

This scenario represents a typical academic performance distribution.

4. Graphical Representation

The graph of student exam scores forms a **bell-shaped curve**:

- The peak of the curve is at 70.
- Scores gradually decrease as we move away from 70.

- The distribution is symmetrical.

The horizontal axis represents exam scores, and the vertical axis represents probability density.

5. The Empirical Rule (68–95–99.7 Rule)

The Empirical Rule applies specifically to normal distributions and states that:

- **68% of observations lie within 1 standard deviation of the mean**
- **95% lie within 2 standard deviations**
- **99.7% lie within 3 standard deviations**

This rule helps us understand how data is distributed without complex calculations.

6. Application of Empirical Rule to Exam Scores

Given:

- $\mu = 70$
- $\sigma = 10$

6.1 Within 1 Standard Deviation ($\mu \pm 1\sigma$)

70−10=6070 - 10 = 60 70−10=60 *70+10=8070 + 10 = 80*70+10=80

68% of students scored between 60 and 80

This means most students performed around average.

6.2 Within 2 Standard Deviations ($\mu \pm 2\sigma$)

70−20=5070 - 20 = 50 70−20=50 *70+20=9070 + 20 = 90*70+20=90

95% of students scored between 50 and 90

Very few students scored below 50 or above 90.

6.3 Within 3 Standard Deviations (μ ± 3σ)

70−30=4070 - 30 = 40 70−30=40 *70+30=10070 + 30 = 100*70+30=100

99.7% of students scored between 40 and 100

Almost all students fall in this range.

7. Interpretation in Educational Context

7.1 Identifying Average Performers

Students scoring between 60 and 80 are considered average performers.

7.2 Identifying High Achievers

Students scoring above 90 are in the top 2.5% approximately.

7.3 Identifying Low Performers

Students scoring below 50 fall into the lower 2.5%.

7.4 Detecting Outliers

If a student scores:

- Below 40
- Above 100 (hypothetically)

They are statistical outliers.

8. Importance of Normal Distribution in Education

1. Helps in grading system design
2. Used in competitive exams evaluation
3. Determines percentile rankings
4. Assists in scholarship selection
5. Supports performance prediction

For example, many entrance exams use normal distribution to standardize scores across different exam sessions.

9. Real-World Relevance Beyond Education

Normal distribution is used in:

- Heights of individuals
- IQ scores
- Measurement errors
- Manufacturing quality control
- Financial risk modeling

This shows its universal importance.

10. Advantages of Using Empirical Rule

- Simple and easy to apply
- No need for complex probability calculations
- Quick estimation of data spread
- Useful for large datasets

11. Limitations

- Applies only to approximately normal data
- Not accurate for skewed distributions
- Real-world data may not perfectly follow bell curve

12. Conclusion

The Normal Distribution is one of the most powerful and widely used statistical models for representing real-world data. In the domain of student exam scores, it provides a clear understanding of how marks are distributed around the average.

Using a mean of 70 and standard deviation of 10, we applied the Empirical Rule to determine:

- 68% of students scored between 60–80
- 95% scored between 50–90
- 99.7% scored between 40–100

This statistical approach helps educators analyze performance, identify outliers, design grading systems, and make data-driven academic decisions.

In conclusion, the combination of Normal Distribution and the Empirical Rule provides a simple yet powerful framework for understanding data spread, predicting outcomes, and making informed decisions in educational and many other domains.