

Assignment-2

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

Normal Distribution and Empirical Rule

Domain Chosen: Human Height Distribution

1. Introduction

In statistics, data collected from real-life situations often follows certain patterns. One of the most important and commonly observed patterns is the **Normal Distribution**, also known as the **Gaussian Distribution** or **Bell Curve**. This distribution is widely used because many natural phenomena tend to follow it, such as human height, weight, blood pressure, measurement errors, and IQ scores. Understanding normal distribution helps researchers analyze data, predict probabilities, and interpret results accurately.

This assignment explains the concept of normal distribution and the empirical rule using a real-world example from the domain of **human height**. Human height is an ideal example because it naturally follows a symmetric pattern where most individuals have average height, while fewer individuals are extremely short or extremely tall. The study of such data helps statisticians understand population characteristics and variability.

2. Domain Selection: Human Height

Human height is one of the best examples of a normally distributed variable. When we measure the height of a large population, we observe that:

- Most people have average height.
- Few people are very short.
- Few people are very tall.

This pattern forms a bell-shaped curve when plotted on a graph. The distribution is symmetrical, meaning the number of people shorter than average is roughly equal to the number of people taller than average.

For this assignment, we assume the following statistical values:

- Mean height = **170 cm**
- Standard deviation = **10 cm**

These values will be used to demonstrate the empirical rule and explain the distribution.

3. What is Normal Distribution?

A **normal distribution** is a probability distribution that is symmetric about its mean. It shows how data values are spread around an average value. The highest point of the curve represents the most common value, and the curve gradually decreases on both sides.

Characteristics of Normal Distribution

- Symmetrical shape
- Mean = Median = Mode
- Bell-shaped curve
- Equal spread on both sides
- Total area under curve = 1

The symmetry indicates that data is evenly distributed around the average value.

4. Graph Explanation (Conceptual)

If we plot a graph of human heights:

- The horizontal axis represents height values.
- The vertical axis represents the number of people.

The graph would look like a bell curve:

- Center → average height (170 cm)
- Left side → shorter individuals
- Right side → taller individuals
- Extreme ends → very rare heights

The peak of the curve shows the height range where most people fall.

5. Mean and Standard Deviation

Two statistical measures define the shape of a normal distribution.

Mean (μ)

The mean is the average value of the dataset. It represents the center of the distribution.

Example:

Average height = 170 cm

Standard Deviation (σ)

Standard deviation measures how much values differ from the mean.

- Small $\sigma \rightarrow$ values close to mean
- Large $\sigma \rightarrow$ values spread out

Example:

Standard deviation = 10 cm

This means most people's height is close to 170 cm.

6. Empirical Rule (68–95–99.7 Rule)

The **Empirical Rule** explains how data is distributed around the mean in a normal distribution.

Rule Summary

- **68%** of values lie within **1 standard deviation**
- **95%** lie within **2 standard deviations**
- **99.7%** lie within **3 standard deviations**

7. Applying Empirical Rule to Height

Given:

Mean = 170 cm

Standard deviation = 10 cm

Within 1 Standard Deviation

$170 \pm 10 \rightarrow$ **160 cm to 180 cm**

About **68% of people** fall in this height range.

Within 2 Standard Deviations

$170 \pm 20 \rightarrow$ **150 cm to 190 cm**

About **95% of people** fall in this range.

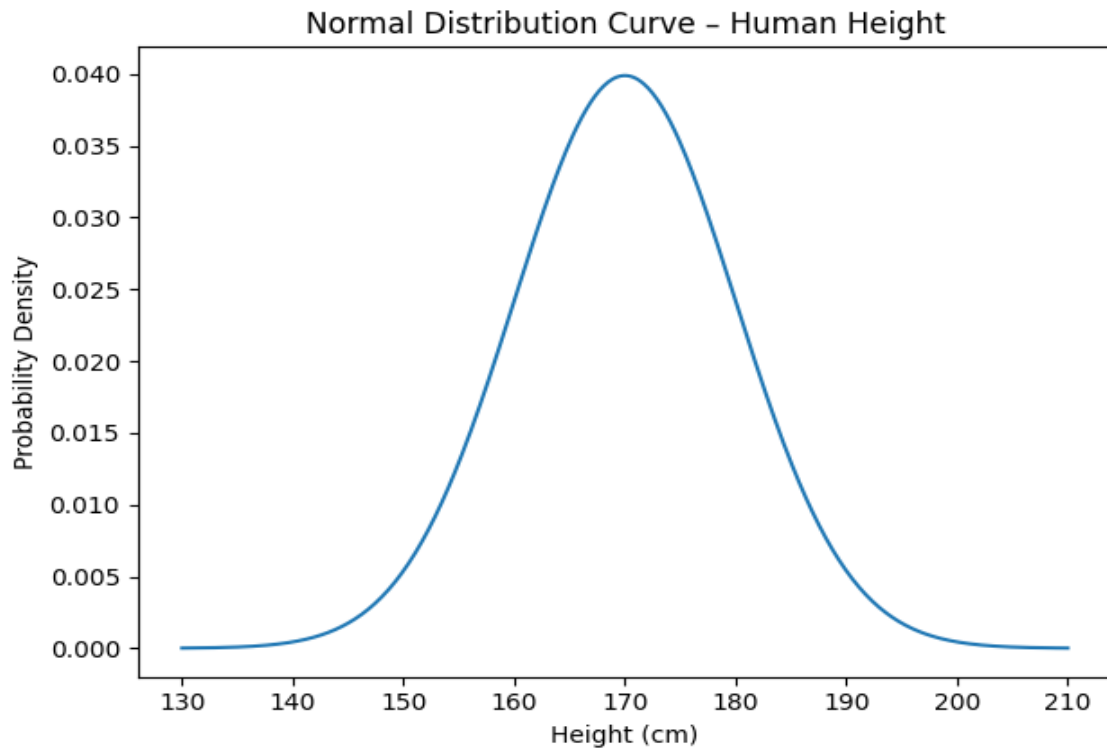
Within 3 Standard Deviations

$170 \pm 30 \rightarrow$ **140 cm to 200 cm**

About **99.7% of people** fall in this range.

Interpretation

- Most people are between 160 and 180 cm tall.
- Almost everyone is between 150 and 190 cm.
- Extremely short or extremely tall individuals are rare.



8. Importance of Normal Distribution

Normal distribution is important because it:

- Helps analyze population characteristics
- Predicts probability of values
- Identifies unusual observations
- Simplifies statistical calculations
- Supports research analysis
- Helps in decision making

Many statistical methods assume data follows a normal distribution.

9. Real-Life Applications

Normal distribution is used in many real-world fields:

Medicine

Analyzing body measurements and health indicators.

Manufacturing

Measuring product defects and quality control.

Psychology

Measuring intelligence scores and behavior patterns.

Economics

Analyzing income distribution and financial data.

Science

Measuring experimental errors.

10. Advantages of Normal Distribution

- Easy to understand visually
- Predictable pattern
- Useful for probability estimation
- Works for many natural datasets
- Simplifies statistical analysis

11. Limitations

- Not all data is normally distributed
- Some data is skewed
- Extreme values may distort results
- Real-world data may deviate slightly

12. Identifying Outliers

Values outside three standard deviations are considered rare or unusual.

In height example:

- Below 140 cm or above 200 cm → very rare heights.

Such values may represent exceptional individuals or measurement errors.

13. Practical Importance

Understanding height distribution helps:

- Clothing manufacturers design sizes
- Doctors evaluate growth patterns
- Sports trainers select athletes
- Researchers study population health

Thus, normal distribution has practical significance in many industries.

14. Summary of Key Points

- Normal distribution is bell-shaped and symmetric.
- Mean represents center of data.
- Standard deviation measures spread.
- Most values lie near the mean.
- Empirical rule explains data percentages.

15. Conclusion

The normal distribution is one of the most fundamental concepts in statistics because it accurately represents how many real-world variables behave. Using human height as an example demonstrates how data naturally clusters around an average, with fewer values appearing as we move away from the center. This bell-shaped pattern allows statisticians and researchers to predict probabilities, identify unusual values, and analyze data efficiently.

The empirical rule further strengthens understanding by showing that most values fall within one, two, and three standard deviations from the mean. In the case of human height, this rule clearly illustrates that average heights are common while extremely short or tall heights are rare. This predictable pattern makes the normal distribution an essential tool in scientific research, data analysis, and real-world decision-making.

Understanding normal distribution and the empirical rule is therefore crucial for students and professionals in fields such as statistics, data science, engineering, medicine, and social sciences. Mastery of these concepts enables accurate interpretation of data, better predictions, and more informed conclusions in practical situations.