

# Skewness - Measuring Asymmetry

## **Detailed Study Notes**

Skewness is a measure of the asymmetry of a probability distribution. It tells you whether data is concentrated on one side of the mean with a tail extending to the other side.

Definition:

Skewness quantifies the degree and direction of asymmetry in a distribution.

Formula:

$$\text{Skewness} = [\sum(x - \mu)^3 / n] / \sigma^3$$

where  $\mu$  is the mean and  $\sigma$  is the standard deviation

## **Types of Skewness:**

1. Zero Skewness ( $\approx 0$ ):

- Symmetric distribution
- Mean  $\approx$  Median  $\approx$  Mode
- Example: Normal distribution

2. Positive Skewness ( $> 0$ ):

- Right-skewed or right-tailed
- Long tail on the right
- Mode  $<$  Median  $<$  Mean
- Example: Income distribution

3. Negative Skewness ( $< 0$ ):

- Left-skewed or left-tailed
- Long tail on the left
- Mean  $<$  Median  $<$  Mode
- Example: Age at retirement

Detailed Example:

Dataset: 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 15

Step 1: Calculate mean

$$\text{Mean} = (1+2+2+3+3+3+4+4+4+5+15) / 10 = 42/10 = 4.2$$

Step 2: Calculate standard deviation  
 $SD \approx 3.88$

Step 3: Calculate skewness  
The value of 15 is an outlier on the right  
Skewness  $\approx 1.87$  (positive)

Interpretation:  
This dataset is right-skewed due to the outlier at 15.

Rules of Thumb:

- $|\text{Skewness}| < 0.5$ : Fairly symmetric
- $0.5 < |\text{Skewness}| < 1$ : Moderately skewed
- $|\text{Skewness}| > 1$ : Highly skewed

## Real-World Examples:

Positive Skewness:

- Income: Most people earn moderate income, few earn very high
- House prices: Many affordable homes, few extremely expensive
- Insurance claims: Most claims are small, few are catastrophic
- Response times: Most quick, few very slow

Negative Skewness:

- Age at death: Most people live to old age
- Test scores with ceiling effect: Most score high
- Product lifetimes with early failures removed

Impact on Analysis:

- Mean is pulled toward the tail
- Median is more representative of typical value
- May need transformation (log, square root)
- Affects choice of statistical tests

Applications:

- Finance: assessing return distributions
- Quality control: identifying process shifts
- Risk management: understanding tail risk
- Market research: analyzing customer behavior