

# Mean, Median, and Mode – Formulae

## 1. Discrete (Ungrouped) Data

Let the observations be

$$x_1, x_2, \dots, x_n$$

### Mean

$$\bar{x} = \frac{\sum x}{n}$$

### Median

Arrange the data in ascending order.

- If  $n$  is odd:

$$\text{Median} = \text{value of } \left( \frac{n+1}{2} \right)^{\text{th}} \text{ item}$$

- If  $n$  is even:

$$\text{Median} = \frac{\left( \frac{n}{2} \right)^{\text{th}} \text{ item} + \left( \frac{n}{2} + 1 \right)^{\text{th}} \text{ item}}{2}$$

### Mode

Mode = value with the highest frequency

## 2. Discrete Grouped Data

(Data given as values  $x_i$  with frequencies  $f_i$ )

### Mean

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

### **Assumed Mean Method**

$$\bar{x} = a + \frac{\sum f_i d_i}{\sum f_i}$$

where

$$d_i = x_i - a$$

and  $a$  is the assumed mean.

### **Median**

$$\text{Median} = \text{value whose cumulative frequency } \geq \frac{N}{2}$$

where

$$N = \sum f_i$$

### **Mode**

$$\text{Mode} = \text{value with maximum frequency}$$

## **3. Continuous Grouped Data**

(Data in class intervals with frequencies)

### **Mean**

$$\bar{x} = \frac{\sum f_i m_i}{\sum f_i}$$

where

$$m_i = \frac{\text{lower limit} + \text{upper limit}}{2}$$

### **Step Deviation Method**

$$\bar{x} = a + \frac{\sum f_i d_i}{\sum f_i} \times h$$

where

$$d_i = \frac{m_i - a}{h}$$

- $h$  = class width
- $a$  = assumed mean

## Median

$$\text{Median} = l + \left( \frac{\frac{N}{2} - cf}{f} \right) \times h$$

where

- $l$  = lower boundary of median class
- $N = \sum f$
- $cf$  = cumulative frequency before median class
- $f$  = frequency of median class
- $h$  = class width

## Mode

$$\text{Mode} = l + \left( \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) \times h$$

where

- $l$  = lower boundary of modal class
- $f_1$  = frequency of modal class
- $f_0$  = frequency of class before modal class
- $f_2$  = frequency of class after modal class
- $h$  = class width

## Empirical Relation (Optional)

$$\text{Mode} = 3(\text{Median}) - 2(\text{Mean})$$

## 2. Discrete (Ungrouped) Data

Let the observations be

$$x_1, x_2, \dots, x_n$$

## Variance

$$\sigma^2 = \frac{1}{n} \sum (x - \bar{x})^2$$

## Standard Deviation

$$\sigma = \sqrt{\frac{1}{n} \sum (x - \bar{x})^2}$$

## Coefficient of Variation

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

## Discrete Grouped Data

(Data given as values  $x_i$  with frequencies  $f_i$ )

### Variance

$$\sigma^2 = \frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i}$$

### Standard Deviation

$$\sigma = \sqrt{\frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i}}$$

### Assumed Mean Method

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{\sum f_i} - \left( \frac{\sum f_i d_i}{\sum f_i} \right)^2}$$

where

$$d_i = x_i - a$$

and  $a$  is the assumed mean.

### Coefficient of Variation

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

## Continuous Grouped Data

(Data given in class intervals with frequencies)

### Variance

$$\sigma^2 = \frac{\sum f_i(m_i - \bar{x})^2}{\sum f_i}$$

## Standard Deviation

$$\sigma = \sqrt{\frac{\sum f_i(m_i - \bar{x})^2}{\sum f_i}}$$

where

$$m_i = \frac{\text{lower limit} + \text{upper limit}}{2}$$

## Step Deviation Method

$$\sigma = h \sqrt{\frac{\sum f_i d_i^2}{\sum f_i} - \left( \frac{\sum f_i d_i}{\sum f_i} \right)^2}$$

where

$$d_i = \frac{m_i - a}{h}$$

- $h$  = class width
- $a$  = assumed mean

## Coefficient of Variation

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

## Bayes' Theorem

If  $A$  and  $B$  are events with  $P(B) \neq 0$ , then

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}$$

If  $A_1, A_2, \dots, A_n$  are mutually exclusive and exhaustive events, then

$$P(A_i | B) = \frac{P(B | A_i) P(A_i)}{\sum_{j=1}^n P(B | A_j) P(A_j)}$$

## Multiplication Theorem of Probability

For two events  $A$  and  $B$  with  $P(B) \neq 0$ ,

$$P(A \cap B) = P(A) P(B | A)$$

Similarly,

$$P(A \cap B) = P(B) P(A | B)$$

## For Three Events

$$P(A \cap B \cap C) = P(A) P(B | A) P(C | A \cap B)$$

If  $A, B, C$  are independent events, then

$$P(A \cap B \cap C) = P(A) P(B) P(C)$$

—

## Inclusion–Exclusion Principle

### For Two Events

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

### For Three Events

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(B \cap C) - P(C \cap A) \\ &\quad + P(A \cap B \cap C) \end{aligned}$$

### General Form (n Events)

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum P(A_i) - \sum P(A_i \cap A_j) + \sum P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} P(A_1 \cap \dots \cap A_n)$$

## Probability Mass Function (PMF)

Let  $X$  be a discrete random variable.

The probability mass function is defined as

$$P(X = x) = p(x)$$

Properties:

- $p(x) \geq 0$  for all  $x$
- $\sum p(x) = 1$

## Cumulative Distribution Function (CDF)

The cumulative distribution function of a random variable  $X$  is

$$F(x) = P(X \leq x)$$

## For Discrete Random Variables

$$F(x) = \sum_{t \leq x} P(X = t)$$

## For Continuous Random Variables

$$F(x) = \int_{-\infty}^x f(t) dt$$

Properties:

- $0 \leq F(x) \leq 1$
- $F(-\infty) = 0$
- $F(\infty) = 1$

## Relation Between PDF and CDF

If  $F(x)$  is differentiable, then

$$f(x) = \frac{d}{dx} F(x)$$

## Mathematical Expectation, Variance and Standard Deviation

Let  $X$  be a random variable.

### Mathematical Expectation (Mean)

#### For a Discrete Random Variable

$$E(X) = \sum x P(X = x)$$

#### For a Continuous Random Variable

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

—

### Variance of a Random Variable

#### Definition

$$\text{Var}(X) = E[(X - \mu)^2]$$

where  $\mu = E(X)$ .

## Alternate Formula

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

## Discrete Case

$$\text{Var}(X) = \sum (x - \mu)^2 P(X = x)$$

## Continuous Case

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

—

## Standard Deviation of a Random Variable

$$\text{SD}(X) = \sqrt{\text{Var}(X)}$$

# Bernoulli Trials and Binomial Distribution

## Bernoulli Trial

A Bernoulli trial is a random experiment with only two possible outcomes:

- Success (with probability  $p$ )
- Failure (with probability  $q = 1 - p$ )

Each trial is independent and the probability of success remains constant.

—

## Binomial Distribution

If a random variable  $X$  denotes the number of successes in  $n$  independent Bernoulli trials, then  $X$  follows a binomial distribution.

## Probability Mass Function

$$P(X = k) = f(k; n, p) = \binom{n}{k} p^k q^{n-k}$$

where

- $k = 0, 1, 2, \dots, n$
- $p$  = probability of success
- $q = 1 - p$

—

## Mean, Variance and Standard Deviation

### Mean

$$E(X) = np$$

### Variance

$$\text{Var}(X) = npq$$

### Standard Deviation

$$\text{SD}(X) = \sqrt{npq}$$

where  $q = 1 - p$ .