# Random Variables, Expectation, and Variance: Core Statistical Concepts

---

## Random Variables

In statistics, a **random variable** is a numerical outcome of a random process or experiment. It assigns a numerical value to each possible outcome of a probabilistic event.

### Types of Random Variables:

1. **Discrete Random Variable**
    - Takes on a **finite or countable** number of distinct values.
    - Example: The number of heads in 3 coin tosses (values: 0, 1, 2, 3).
    - Common Distributions: Binomial, Poisson, Geometric.
2. **Continuous Random Variable**
    - Can take an **infinite number of values** within a range.
    - Example: The height of students in a class (values can be any real number within a range).
    - Common Distributions: Normal, Exponential, Uniform.

### Notation:

- A random variable is typically denoted by **X, Y, or Z**.
- A specific value of a random variable is written in lowercase, like **x or y**.
- Probability of a discrete random variable taking a value:

  $$P(X = x)$$

- Probability for a continuous random variable is given by a probability density function (PDF), where probabilities are found using integration.

---

# Expected Value (Mean of a Random Variable)

- **Expected Value ($E(X)$):** The **long-run average** of X.

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

  - **Interpretation:** Theoretical mean over many trials.
  - **Example:** Rolling a fair die, X={1,2,3,4,5,6}, $E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$

🔷 **Key Point: Expected value is a weighted average of possible values.**

## Properties of Expected Values

- **Constant Rule**: $E(c) = c$ (A constant's expected value is itself).
- **Linearity**:

$$E(aX + b) = aE(X) + b$$

- **Sum Rule**:

$$E(X + Y) = E(X) + E(Y)$$

(Valid even if X and Y are dependent).

---

# Distribution Functions in Statistics

A **distribution function** describes the probabilities of a random variable taking certain values. There are two main types

## Cumulative Distribution Function (CDF)

The **CDF** of a random variable $X$, denoted as $F(x)$, gives the probability that $X$ takes a value less than or equal to $x$:

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

- For **discrete random variables**, it sums up probabilities up to $x$.
- For **continuous random variables**, it is found by integrating the probability density function (PDF).

## Probability Density Function (PDF) [For Continuous Variables]

The **PDF**, denoted as $f(x)$, represents the likelihood of a random variable taking a specific value.

$P(a \leq X \leq b) = \int_a^b f(x)dx$

- The area under the curve of the PDF over an interval gives probability.
- Example: The **Normal Distribution** has the famous bell-shaped PDF.

## Probability Mass Function (PMF) [For Discrete Variables]

The **PMF**, denoted as $P(X = x)$, gives the probability of a discrete random variable taking exact values.

$$P(X = x) = \text{some probability value}$$

- Example: A fair die roll has

$$P(X = 1) = P(X = 2) = \cdots = P(X = 6) = \frac{1}{6}$$

## Relation between PDF and CDF

**1. Probability Density Function (PDF), denoted as f(x):**

- **What it is:** The PDF describes the relative likelihood of a continuous random variable taking on a given value. Think of it as the "shape" of the distribution.
- **Key property:** The area under the PDF curve within a certain interval represents the probability that the random variable falls within that interval.
- **Important note:** The PDF itself doesn't directly give you probabilities. You need to integrate it to find probabilities.

**2. Cumulative Distribution Function (CDF), denoted as F(a):**

- **What it is:** The CDF gives the probability that a random variable X takes on a value less than or equal to a specific value 'a'.
- **In other words:** It accumulates the probabilities up to a certain point.
- **Always increasing:** As 'a' increases, the CDF will either stay the same or increase (it can't decrease).
- **Range:** The CDF ranges from 0 to 1. F(-∞) = 0 and F(∞) = 1.

**3. The Relationship: $F(a) = P\{X \in (-\infty, a]\} = \int_{-\infty}^{a} f(x)\,dx$**

> ⚠️ **(-∞, a]:** This notation means the interval that starts from negative infinity (meaning it includes all values below 'a') and goes up to and *includes* 'a'. **The square bracket ']' indicates that 'a' is included in the interval.**

This is the core of the explanation:

- **F(a):** This is the CDF evaluated at 'a'.
- **P{X ∈ (-∞, a]}:** This reads "the probability that the random variable X is in the interval from negative infinity up to and including 'a'".
- **∫(-∞ to a) f(x) dx:** This is the integral of the PDF f(x) from negative infinity up to 'a'.

**In simple terms:**

The CDF at a point 'a' is equal to the area under the PDF curve from negative infinity up to 'a'.

**Why is this important?**

- **Connecting concepts:** It shows how the PDF and CDF are intrinsically linked.
- **Calculating probabilities:** It provides a way to calculate probabilities using the PDF, which is often easier to derive or know.
- **Understanding distributions:** It helps in understanding the overall behavior and characteristics of a probability distribution.

**Example:**

Imagine you have a normal distribution (bell curve). The PDF would show the shape of the bell, with the peak representing the most likely value. The CDF at a point 'a' would show the area under the bell curve to the left of 'a', representing the probability of getting a value less than or equal to 'a'.

**In essence, the CDF "accumulates" the probability information provided by the PDF.**

---

# Jointly Distributed Random Variables

When **two or more random variables** are considered together, and their probabilities depend on each other, they are called **jointly distributed random variables**.

# 1. Understanding Joint Distribution

- Suppose we have **two random variables**, $X$ and $Y$, that represent two different measurements from the same experiment.
- Their **joint distribution** describes the probability of different combinations of $X$ and $Y$ occurring together.

For example:

- Let $X$ = number of heads in two coin tosses.
- Let $Y$ = number of tails in two coin tosses.
- Since they come from the same experiment, X and Y are **jointly distributed**.

# 2. Types of Joint Distributions

## (a) Joint Probability Mass Function (PMF) – Discrete Case

- If X and Y are discrete random variables, their joint probability mass function is:

$$P(X = x, Y = y) = P(X, Y)$$

- Example: Rolling a die twice
  - $X$ = result of first roll, $Y$ = result of second roll.
  - $P(X = 2, Y = 5)$= probability of rolling a **2 first, then a 5**.

## (b) Joint Probability Density Function (PDF) – Continuous Case

- If $X$ and $Y$ are continuous, their joint probability density function is:

$$f(x, y) \geq 0, \quad \text{and} \quad \iint f(x, y)\, dx\, dy = 1$$

- The probability of $X$ and $Y$ being in a certain region is:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y)\, dx\, dy$$

- Example: Heights of people ($X$) and weights of people ($Y$).

---

# Independent Random Variables

Two random variables $X$ **and** $Y$ are **independent** if knowing the value of one **does not provide any information** about the other. In other words, their probabilities are completely separate.

## Definition of Independence

### (a) Discrete Random Variables

For discrete random variables, $X$ and $Y$ are independent if their **joint probability mass function (PMF)** satisfies:

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y) \quad \text{for all } x, y.$$

This means that the probability of two events happening together is just the product of their individual probabilities.

### (b) Continuous Random Variables

For continuous random variables, $X$ and $Y$ are independent if their **joint probability density function (PDF)** satisfies:

$$f(x, y) = f_X(x) \cdot f_Y(y) \quad \text{for all } x, y.$$

This means that the joint PDF is simply the product of the individual (marginal) PDFs.

### Intuition Behind Independence

If two random variables are independent:

✅ Knowing $X$ tells us **nothing** about $Y$.

✅ Their probabilities behave as if they come from separate experiments.

📌 **Example (Discrete Case):**

- Let X = result of rolling a **fair 6-sided die**.
- Let Y = result of flipping a **fair coin (1 = heads, 0 = tails)**.
- Since rolling a die has no effect on flipping a coin, the probabilities remain separate:

$P(X=3, Y=1) = P(X=3) \cdot P(Y=1) = (\frac{1}{6} \times \frac{1}{2}) = \frac{1}{12}$

So $X$ and $Y$ are **independent.**

**Example (Continuous Case):**

- Suppose X and Y are independent normal variables:

  $$X \sim N(0,1), \quad Y \sim N(5,2)$$

- Then their **joint PDF** is simply:

  $$f(x,y) = f_X(x) \cdot f_Y(y)$$

- This confirms independence because their probabilities do not influence each other.

---

# Expectation in Statistics (Expected Value, E[X])

The **expectation (expected value)** of a random variable is a measure of its **long-run average** or **mean** when an experiment is repeated many times. It represents the **center of the probability distribution**.

## Definition of Expectation

The expectation (denoted as $E[X]$ or $\mathbb{E}[X]$) is the **weighted average of all possible values of $X$**, where the weights are their probabilities.

### (a) Expectation for a Discrete Random Variable

If X is a **discrete** random variable with values $x_1, x_2, \ldots, x_n$ and probabilities $P(X = x_i)$, then:

$$E[X] = \sum_i x_i P(X = x_i)$$

📌 **Example:**

Suppose a dice roll gives X values **1, 2, 3, 4, 5, and 6**, each with probability $\frac{1}{6}$.

The expectation is:

$$E[X] = (1 \times \frac{1}{6}) + (2 \times \frac{1}{6}) + (3 \times \frac{1}{6}) + (4 \times \frac{1}{6}) + (5 \times \frac{1}{6}) + (6 \times \frac{1}{6}) = 3.5$$

### (b) Expectation for a Continuous Random Variable

If $X$ is **continuous** with probability density function (PDF) $f(x)$, then the expectation is given by:

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

📌 **Example:**

For a uniform distribution $X \sim U(0, 1)$, where $f(x) = 1$ for $0 \leq x \leq 1$, we compute:

$E[X] = \int_0^1 x \cdot 1 \, dx = \frac{1}{2}$

## Properties of Expectation

1. **Linearity of Expectation:**
    - For any two random variables X and Y,

      $$E[aX + bY] = aE[X] + bE[Y]$$
      (where a,b are constants).

    - Example: If $E[X] = 3, E[Y] = 5$, then:

      $$E[2X + 3Y] = 2(3) + 3(5) = 6 + 15 = 21$$

2. **Expectation of a Constant:**$E[c] = c$

   (The expectation of a constant is just the constant itself.)

3. **Expectation of a Sum:**
    - If $X_1, X_2, ..., X_n$ are random variables:

      $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$

4. **If $X$ and $Y$are Independent:**$E[XY] = E[X]E[Y]$

   But if $X$ and $Y$ are **not** independent, this does **not** necessarily hold.

---

# Variance

**Variance** is a measure of how much a set of numbers (or a random variable) deviates from its **mean (expected value)**. It tells us how **spread out** or **dispersed** the data is.

# Definition of Variance

The variance of a random variable X, denoted as Var(X) or $\sigma^2$, is the **expected squared deviation** from the mean:

$$Var(X) = E[(X - E[X])^2]$$

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - E[2\mu X] + E[\mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - \mu^2 \end{aligned}$$

$$E[X^2] = Var(X) + \mu^2 \qquad \mu = E[X] \text{ (Since E[X] is the mean)}$$

$$E[X^2] = Var(X) + (E[X])^2$$

This means we:

1. **Find the mean** E[X] (expected value).
2. **Calculate each value's deviation from the mean** $(X - E[X])$.
3. **Square each deviation** (to make all values positive).
4. **Take the expected value** (average) of the squared deviations

# Variance Formula

Depending on whether X is **discrete** or **continuous**, we compute variance differently.

## (a) Variance for Discrete Random Variables

For a discrete random variable X with values $x_1, x_2, \ldots, x_n$ and probabilities $P(X = x_i)$, the variance is:

$$Var(X) = \sum_i P(X = x_i)(x_i - E[X])^2$$

📌 **Example (Rolling a Fair Die)**:

- X = outcome of rolling a fair 6-sided die.
- Possible values: 1, 2, 3, 4, 5, 6.
- Each value has equal probability $\frac{1}{6}$.

- The mean (expected value) is: $E[X] = \sum x_i P(X = x_i) = \frac{1+2+3+4+5+6}{6} = 3.5$
- Now, compute variance:

  $Var(X) = \sum_{i=1}^{6} P(X = x_i)(x_i - 3.5)^2$
- After solving, we get:

  $Var(X) = \frac{35}{12} \approx 2.92$

### (b) Variance for Continuous Random Variables

For a **continuous** random variable with probability density function (PDF) $f(x)$, variance is:

$Var(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f(x)\, dx$

📌 **Example (Uniform Distribution U(0,1))**

For $X \sim U(0,1)$, where $f(x) = 1$ for $0 \leq x \leq 1$:

1. Compute $E[X]$:

   $E[X] = \int_0^1 x \cdot 1\, dx = \frac{1}{2}$
2. Compute variance:

   $Var(X) = \int_0^1 (x - \frac{1}{2})^2 \cdot 1\, dx = \frac{1}{12}$

## Properties of Variance

1. **Variance of a Constant:**

   $Var(c) = 0$

   (A constant does not vary.)
2. **Scaling Property:**

   $Var(aX) = a^2 Var(X)$

   (Multiplying by a constant $a$ scales the variance by $a^2$.)
3. **Sum of Independent Random Variables:**

   $Var(X + Y) = Var(X) + Var(Y)$

   If X and Y are independent.
4. **Variance and Standard Deviation Relationship:**

$$\sigma = \sqrt{Var(X)}$$

Standard deviation (SD) is the **square root of variance**, giving a measure of spread in the same units as the data.