**DMV\_U1**

**Random Variable (RV) — The Big Picture**

A **random variable** is like a bridge between a real-world random experiment and numbers we can work with mathematically.

* It **assigns a numerical value** to each possible outcome of a random process.
* Formally, it’s a function from the *sample space* (all possible outcomes) to real numbers.

Example:  
If we roll a die, the random variable XXX could be defined as “the number showing on the die.”  
Possible outcomes → {1, 2, 3, 4, 5, 6}.

**1. Discrete Random Variable (DRV)**

* **Definition:** Can only take **finite** or **countably infinite** values.
* **Key property:** We can list all possible values, even if the list is very long.
* **Probability tool:** **Probability Mass Function (PMF)** P(X=x)P(X = x)P(X=x) → gives the probability of each possible value.
* The sum of all probabilities is **1**.

**Example:** Number of students in a class.

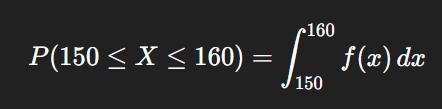
* Possible values: {0, 1, 2, 3, …}
* We can assign:

**Common discrete distributions:** Binomial, Poisson, Geometric.

**2. Continuous Random Variable (CRV)**

* **Definition:** Takes **uncountably infinite** values within an interval or multiple intervals.
* **Key property:** The probability of taking any *exact* single value is **zero**. We talk about **probability over intervals** instead.
* **Probability tool:** **Probability Density Function (PDF)** f(x)f(x)f(x) →
  + The **area under the curve** of the PDF over an interval gives the probability.
  + The total area under the curve = 1.

**Example:** Height of a person.

* Possible values: 150.000… cm, 150.001 cm, 150.002 cm, …
* To find the probability that height is between 150 cm and 160 cm: 

**Common continuous distributions:** Normal, Uniform, Exponential.

## ****Independence for Random Variables****

Just like **two events** are independent if one happening doesn’t affect the other,  
**two random variables** are independent if knowing the value of one gives you **no information** about the value of the other.

**Mathematical Definition**

The mathematical definition of independence depends on the type of random variable.

* Discrete Case: For discrete variables, X and Y are independent if for all possible values x and y:

P(X=x,Y=y)=P(X=x)⋅P(Y=y)

This means the probability of both events occurring together (the joint probability) is simply the product of their individual probabilities (their marginal probabilities).

* Continuous Case: For continuous variables, independence is defined using their probability density functions (PDFs). X and Y are independent if their joint PDF, f,X,Y​(x,y), can be factored into the product of their individual (marginal) PDFs, fX​(x) and fY​(y):

fX,Y​(x,y)=fX​(x)⋅fY​(y)

**Intuition and Examples**

The core idea is that if X and Y are independent, knowing the value of X doesn't change the **probability distribution** of Y, and vice versa.

* **Example 1 (Discrete):** Consider two independent coin tosses. Let X be the result of the first toss (1 for heads, 0 for tails) and Y be the result of the second. The probability of getting heads on the first toss is P(X=1)=0.5, and the probability of getting heads on the second is P(Y=1)=0.5. Since the tosses are independent, the joint probability of getting heads on both is P(X=1,Y=1)=0.25, which is exactly P(X=1)⋅P(Y=1). This confirms their independence.

**Events vs. Random Variables**

It's important to distinguish between these two concepts:

* **Event Independence** focuses on whether the **outcome** of one specific event (e.g., getting a 6 on a die) changes the probability of another specific event.
* **Random Variable Independence** is a more powerful concept that focuses on whether the **entire probability distribution** of one variable remains unchanged regardless of the value of the other.

### Covariance: A Measure of Joint Variability

**Covariance** is a statistical tool that describes how two random variables change in relation to each other. It indicates the **direction** of the linear relationship between them.

### Interpretation of Covariance

* **Positive Covariance (>0)**: When one variable tends to increase, the other also tends to increase. Think of a positive trend.
  + **Example**: As daily temperatures rise (X), ice cream sales also tend to increase (Y).
* **Negative Covariance (<0)**: When one variable increases, the other tends to decrease. This indicates an inverse relationship.
  + **Example**: The more hours a student spends playing video games (X), the lower their exam score tends to be (Y).
* **Zero Covariance (=0)**: There's no consistent linear relationship between the two variables. Their movements are not predictably linked in a linear fashion.

### Mathematical Definition

For two random variables, X and Y, the covariance is defined as the expected value of the product of their deviations from their respective means.

Cov(X,Y)=E[(X−μX​)(Y−μY​)]

Where:

* E[⋅] is the expected value operator.
* μX​=E[X] is the mean of X.
* μY​=E[Y] is the mean of Y.

An alternative, often more practical, formula is:

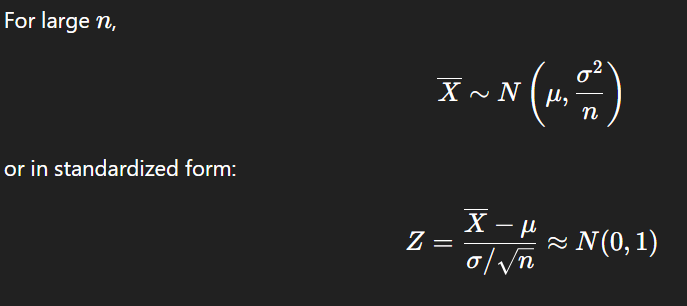
Cov(X,Y)=E[XY]−E[X]E[Y]

### Covariance vs. Independence

It's crucial to remember that **zero covariance does not necessarily imply independence**. While independent variables always have a covariance of zero, a covariance of zero only means there is no **linear** relationship. There could still be a non-linear relationship between the variables.

**The Central Limit Theorem (CLT)** is a statistical principle that states that the **distribution of sample means** will be approximately normal for a large enough sample size, regardless of the original population's distribution. This powerful theorem is the cornerstone of many statistical methods.

### Mathematical Formulation ✍️



### Key Conditions and Implications

The CLT holds true under a few key conditions:

* The samples must be **independent** and **identically distributed** (i.i.d.).
* The population must have a **finite variance**.
* The sample size (n) must be **large enough** (a common rule of thumb is n≥30).

The primary implication of the CLT is that it allows us to use normal distribution-based techniques, such as **confidence intervals** and **hypothesis testing**, even when we don't know the original population's distribution.

**Example Calculation**

Consider a population with a mean of μ=50 and a standard deviation of σ=20. If we take a sample of n=100, we can determine the properties of the sampling distribution of the sample mean.

The **standard error** is:

SE=n​σ​=100​20​=2

According to the CLT, the distribution of our sample means (X) will be approximately normal with a mean of 50 and a standard deviation (standard error) of 2. We can write this as X∼N(50,22). This means we can expect most sample means to fall within a predictable range around 50.

## ****Advantages (Pros)****

1. **Universality:** Works for any population shape if the sample size is large enough.
2. **Foundation for Statistics:** Justifies using normal-based methods for hypothesis testing, confidence intervals, etc.
3. **Simplifies Analysis:** Converts complex or unknown distributions into a predictable normal form.
4. **Practical Application:** Works well in real-world scenarios where exact population distribution is unknown.

## ****Disadvantages (Cons)****

1. **Sample Size Requirement:** Needs a sufficiently large nnn for accuracy, especially for skewed or heavy-tailed data.
2. **Not Always Exact:** For small samples from non-normal populations, the approximation can be poor.
3. **Finite Variance Requirement:** Does not hold for distributions with infinite variance (e.g., Cauchy).
4. **Independence Assumption:** Fails if data points are highly dependent or correlated.