Central Limit Theorem Explained

1. What is the Mean?

The **mean** (or average) of a dataset is the sum of all values divided by the number of values.

Formula:

mean = (x1 + x2 + x3 + ... + xn) / n

It represents the central value of the data.

---

2. What is the Standard Deviation?

The **standard deviation** measures how spread out the numbers are from the mean.

Steps to calculate standard deviation:

1. Find the mean
2. Subtract the mean from each value and square the result
3. Take the average of these squared differences (this is the variance)
4. Take the square root of the variance

Formula:

standard deviation (sigma) = sqrt((1/n) * sum((xi - mean)^2))

A smaller standard deviation means the data points are close to the mean. A larger standard deviation means the data is more spread out.

$P \rightarrow x_1 \quad x_2 \quad x_3 \quad x_4 \cdots$

$\mu \quad \sigma$

3. Central Limit Theorem (CLT)

$\mu_1 \; [x_1 \quad x_2 \ldots x_9]$ 300

$\mu_2 \; [x_3 \quad x_7, x_9]$

The **Central Limit Theorem** states:

> If you take many random samples of size n from any population (with finite mean and variance), then the distribution of the sample means will tend to be **approximately normal** as n becomes large — regardless of the shape of the original population.

---

4. Mathematical Expression

Let X1, X2, ..., Xn be n independent, identically distributed (i.i.d) random variables with:

- Mean = mu (μ)
- Standard deviation = sigma (σ)

Then the **sampling distribution of the sample mean** (denoted as X̄) approaches a normal distribution with:

- Mean = mu (μ)
- Standard deviation = sigma / sqrt(n) $\left(\frac{\sigma}{\sqrt{n}}\right)$

uniform distribution