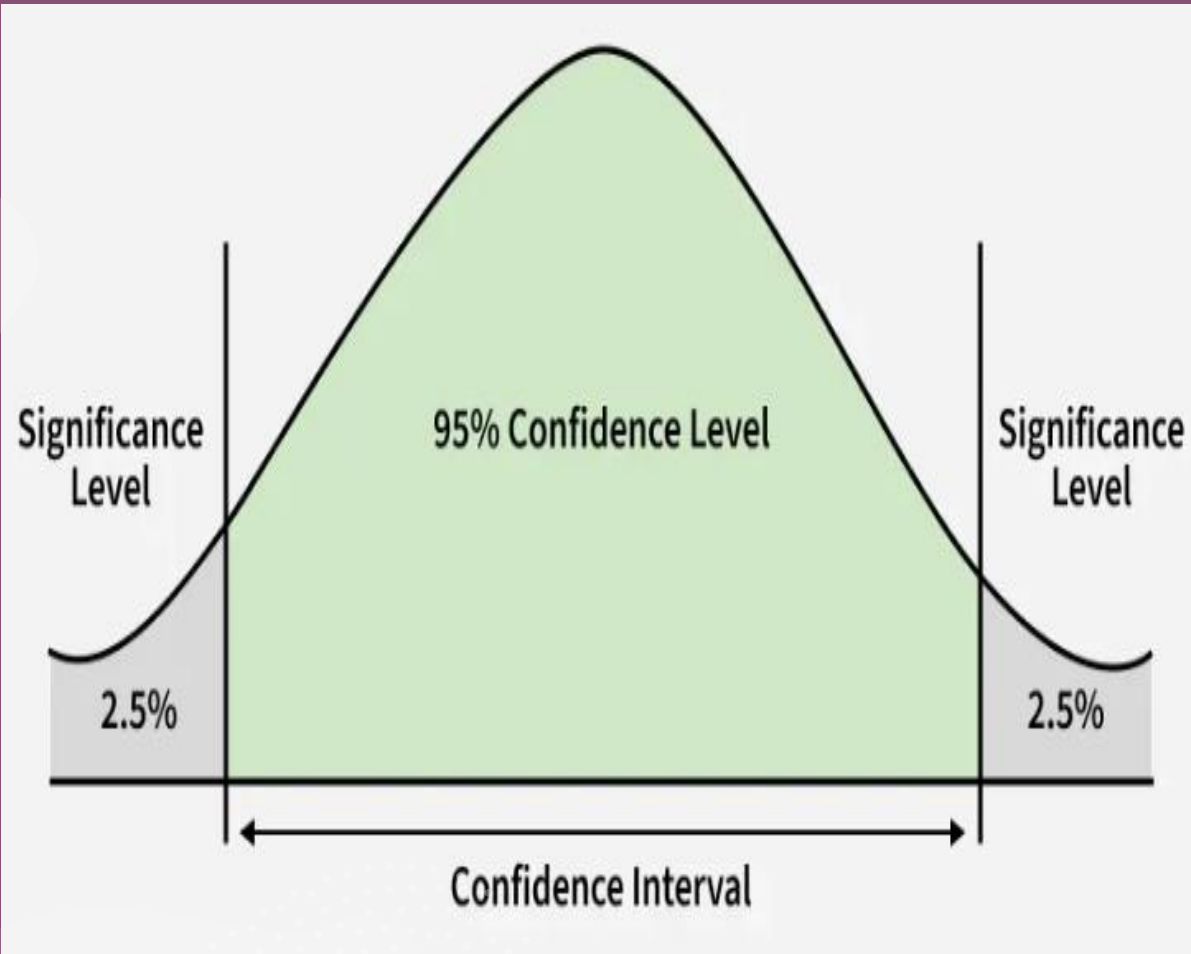


Confidence Interval (CI)

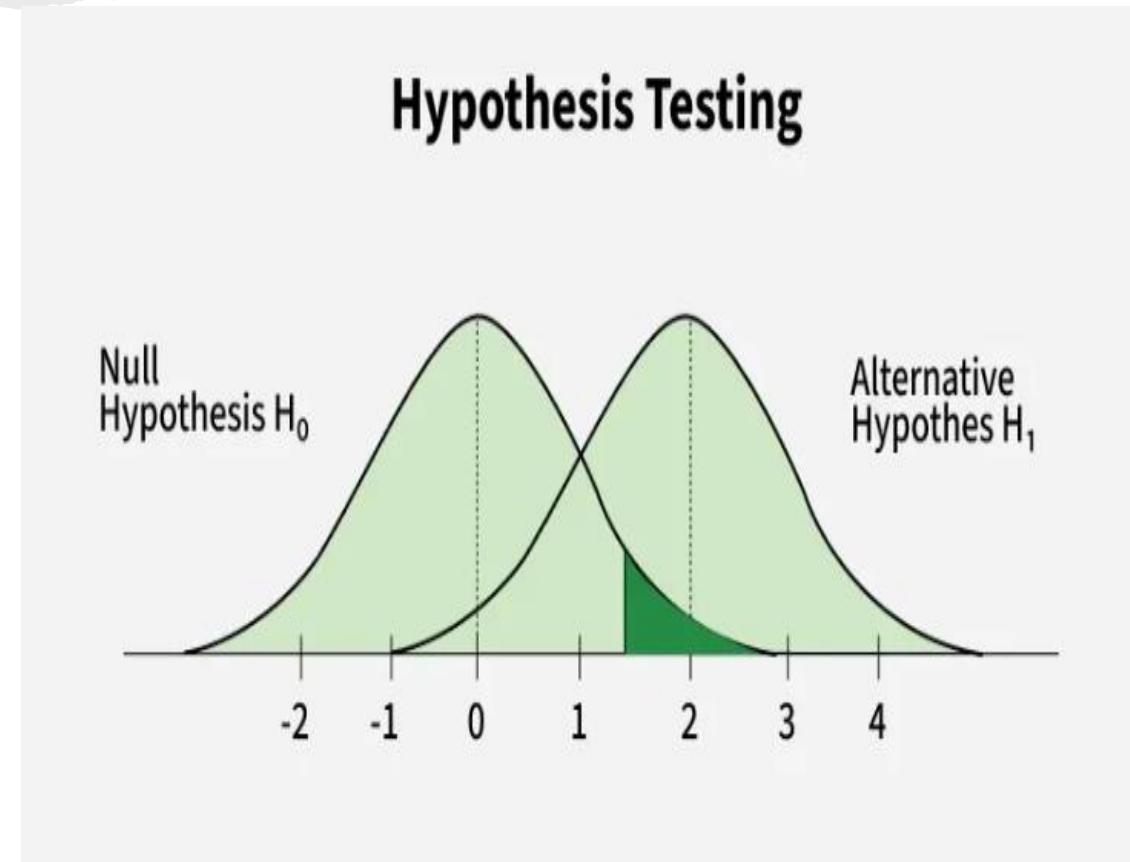


- **Meaning:**
A confidence interval is a range that estimates a population parameter with a given confidence level (e.g., 95%).
- **General Form:**
 $\text{Estimate} \pm (\text{Critical Value} \times \text{Standard Error})$
- **Computation – Different Cases**
- **Critical Values**
- **90%: 1.645 95%: 1.96 99%: 2.576**

Hypothesis Testing

Definition: Statistical method to decide whether sample evidence supports or rejects a claim about a population parameter

- **Null Hypothesis (H_0):** Assumes *no effect / no difference*
- **Alternative Hypothesis (H_1):** Assumes *effect / difference exists*
- **Significance Level (α):** Probability of rejecting H_0 when it is true (commonly 0.05)
- **Test Statistic:** Numerical value from sample data (z , t , χ^2 , F)
- **p-value:** Probability of observing data as extreme as the sample assuming H_0 is true
- **Decision Rule:**
 - If **p-value $\leq \alpha$ \rightarrow Reject H_0**
 - If **p-value $> \alpha$ \rightarrow Fail to reject H_0**
- **Types of Errors:**
 - **Type I Error:** Rejecting true H_0
 - **Type II Error:** Failing to reject false H_0



KS Test

- **Kolmogorov–Smirnov (KS) Test**
- **Definition:** Non-parametric test to compare probability distributions using CDFs
- **Purpose:**
 - One-sample: Compare sample with a theoretical distribution
 - Two-sample: Compare two samples
- **Hypotheses:**
 - H_0 : Distributions are the same
 - H_1 : Distributions are different
- **Test Statistic:**
- $D = \max |F_1(x) - F_2(x)|$ $D = \max |F_1(x) - F_2(x)|$ $D = \max |F_1(x) - F_2(x)|$
- (Maximum vertical distance between CDFs)
- **Decision Rule:**
 - If $p\text{-value} \leq \alpha \rightarrow \text{Reject } H_0$
 - If $p\text{-value} > \alpha \rightarrow \text{Fail to reject } H_0$
- **Key Features:**
 - Distribution-free, works on entire distribution
 - Sensitive to differences in location and shape
- **Assumptions & Limits:**
 - Continuous, independent data
 - Less effective for discrete data or tail differences
- **Applications:**
 - Normality testing, model validation, distribution comparison

One sample Vs Two sample KS test

One-Sample KS Test	Two-Sample KS Test
Employed to assess whether a single sample of data conforms to a specific theoretical distribution.	Utilized to evaluate whether two independent samples originate from the same underlying distribution.
Compares the (EDF) of the sample with the (CDF) of the theoretical distribution.	It compares the EDF of one sample with the EDF of the other sample.
Null hypothesis assumes that the sample is drawn from the specified distribution.	Null hypothesis posits that the two samples are drawn from identical distributions.
Test statistic, represents the maximum vertical deviation between the EDF and CDF.	The test statistic, reflects the maximum difference between the two EDFs.

