T-Test & T-Distribution: A Complete Guide with Diagrams and Interview Q&A

# What is a T-Test?

A t-test is a statistical hypothesis test used to determine whether there is a significant difference between the means of two groups, especially when the sample size is small and population standard deviation is unknown. It is based on the t-distribution.

# When to Use a T-Test vs. Z-Test

- Use a Z-test when the population standard deviation is known and sample size is large (n > 30).  
- Use a T-test when the population standard deviation is unknown or sample size is small (n ≤ 30).

# Types of T-Tests

## 1. One-Sample T-Test

Tests whether the sample mean is significantly different from a known or hypothesized population mean.

## 2. Two-Sample T-Test (Independent)

Compares the means of two independent groups (e.g., test scores of two different classes).

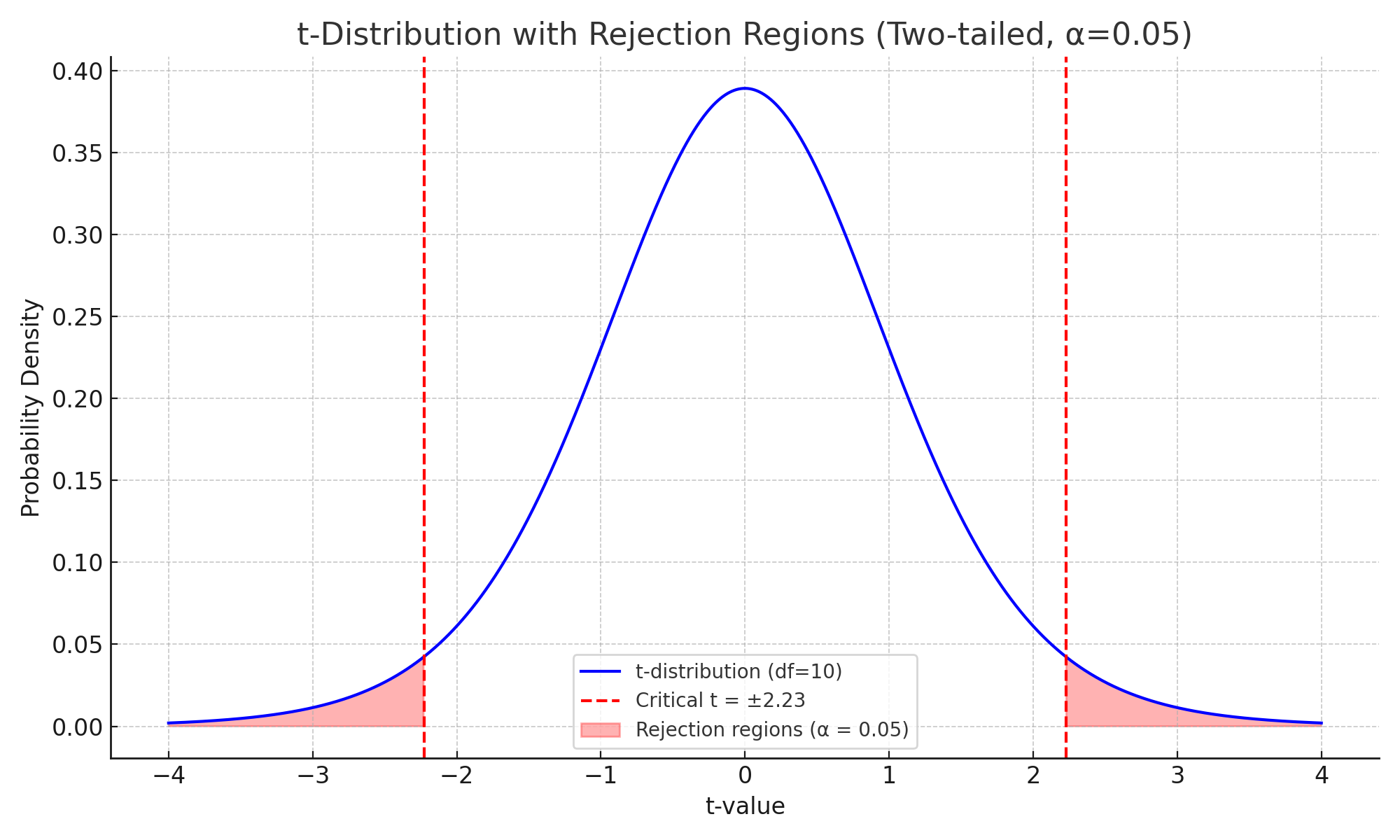
## 3. Paired T-Test

Compares the means from the same group at different times (e.g., before and after treatment).

# What is the T-Distribution?

The t-distribution is a symmetric, bell-shaped distribution like the normal distribution but has heavier tails. It accounts for the added uncertainty from estimating the population standard deviation with a sample standard deviation. As sample size increases, it approaches the normal distribution.

## Critical Region in T-Distribution



The shaded regions represent the areas where the null hypothesis would be rejected at α = 0.05.

# P-Value in a T-Test

In a t-test, the p-value represents the probability of observing the sample data (or something more extreme) assuming the null hypothesis is true. If the p-value is less than the chosen significance level (e.g., 0.05), we reject the null hypothesis.

# Interview-Style Questions

## Q1: When should you prefer a t-test over a z-test in a real project?

A: Prefer a t-test when your sample size is small (n ≤ 30) or the population standard deviation is unknown. For example, when comparing average conversion rates between two website versions with limited data.

## Q2: What is the difference between a paired and two-sample t-test?

A: A paired t-test compares two related samples, like before-and-after measurements on the same subjects. A two-sample t-test compares two independent groups. For instance, compare weight loss in two diet groups (independent) vs. measuring weight loss in the same individuals before and after (paired).

## Q3: How do you interpret a p-value of 0.03 in a t-test?

A: If your significance level is 0.05, a p-value of 0.03 means there's only a 3% chance you'd observe a difference this extreme (or more) assuming the null hypothesis is true. Since 0.03 < 0.05, you reject the null.