



JOHNS HOPKINS

WHITING SCHOOL
of ENGINEERING

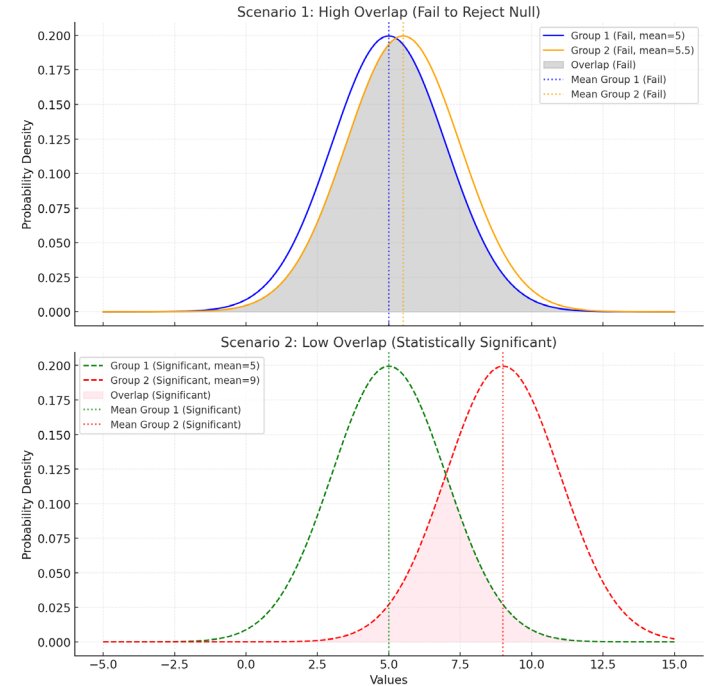
Design of Experiments

ANOVA and Statistical Tests

T-Test

Key Steps

1. Calculate the means
2. Calculate the standard error
3. Compute the degrees of freedom
4. Determine the p-value



T-Test: Mathematical Formulation

- **T-statistic Calculation**

$$t = \frac{\overline{x_1} - \overline{x_2}}{SE}$$

Where:

- $\overline{x_i}$: Mean of sample i
- SE: Standard Error of the Difference
- **SE** Calculation depends on observed variances
- If samples have equal variances: Pooled T-Test
- If not, Welch's T-Test

T-Test: Mathematical Formulation

- **T-statistic Calculation**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE}$$

Where:

- \bar{x}_i : Mean of sample i
- SE: Standard Error of the Difference

- **Pooled T-Test (Equal Variances)**

$$SE = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

- **Welch's T-Test (Unequal Variances)**

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Where:

- s_p : Standard deviation of pooled data
 s_i : Standard deviation of sample i
 n_i : Size of sample i

T-Test: Mathematical Formulation, Degrees of Freedom (df)

- **Pooled T-Test**

$$df = n_1 + n_2 - 2$$

- **Welch's T-Test**

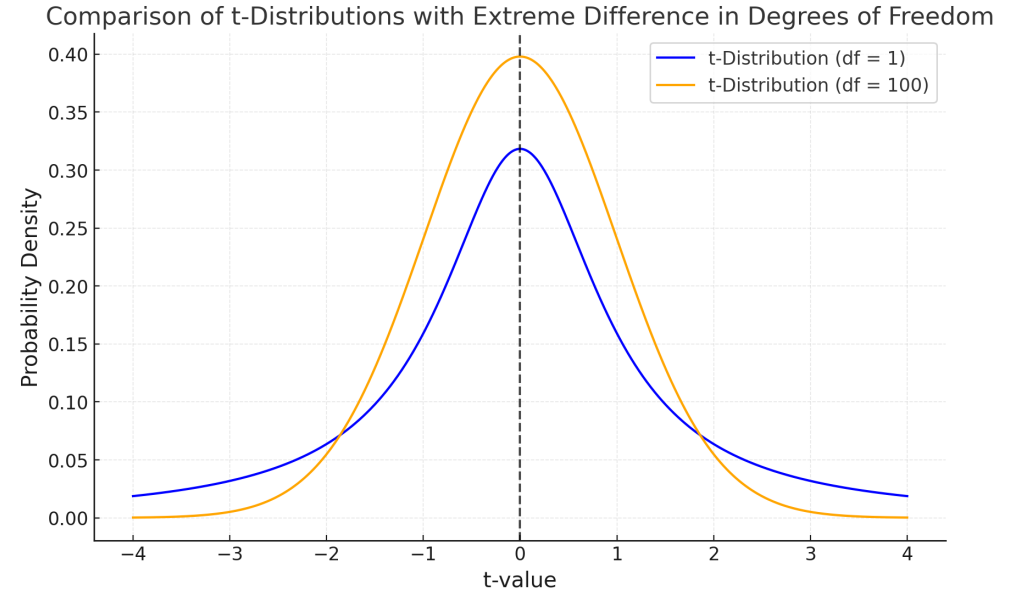
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Where:

s_p : Standard deviation of pooled data

s_i : Standard deviation of sample i

n_i : Size of sample i



T-Test: Mathematical Formulation, p-value

- **1-Tailed Test**

$$p = 1 - F_t(t, df) \text{ if } t > 0$$

$$p = F_t(t, df) \text{ if } t < 0$$

- **2-Tailed Test**

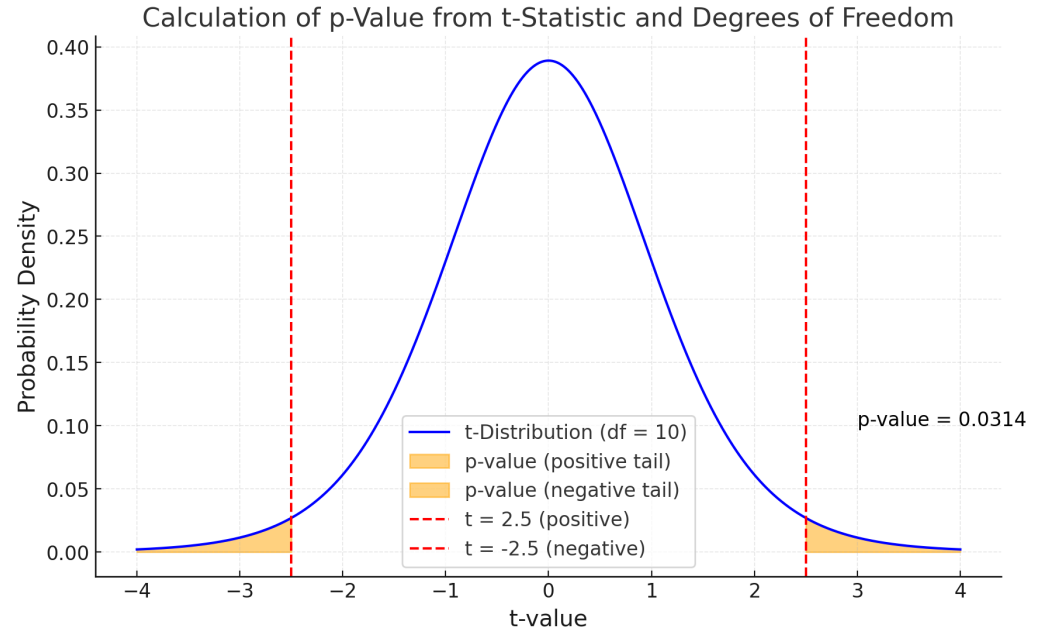
$$p = 2 \cdot (1 - F_t(|t|, df))$$

Where:

t : t-statistic

df : Degrees of freedom

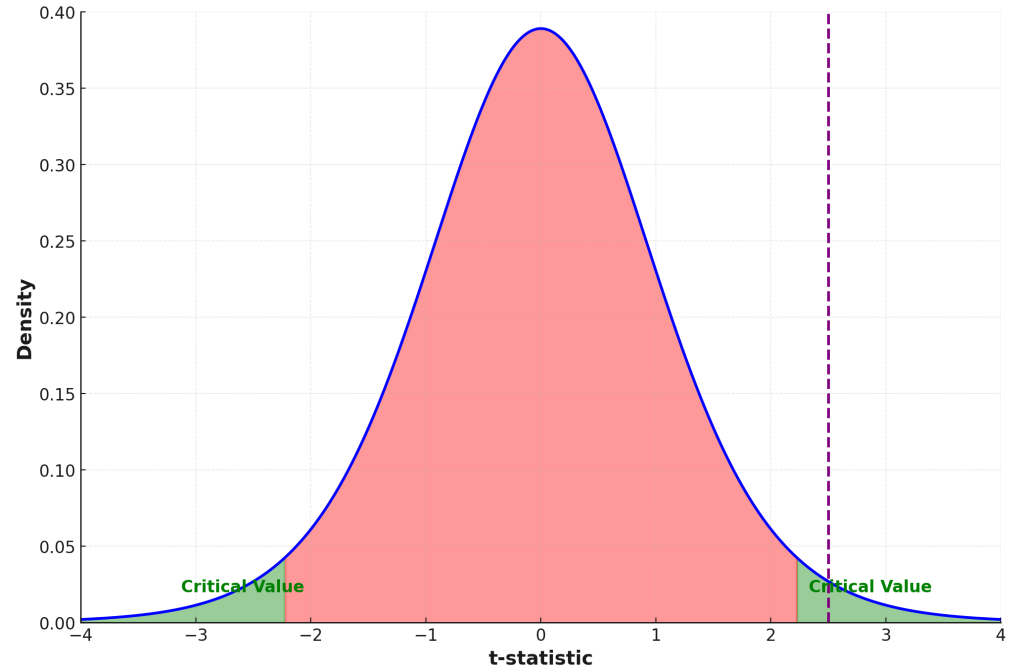
$F_t(t, df)$: Cumulative distribution function (CDF) of the t-distribution given t and df



T-Test: Make a Decision

Reject H_0 if the p-value is less than α ; otherwise, **fail to reject H_0** .

In the example on the right, we would **reject H_0** as our t-statistic (dotted line) is to the right of the critical value



T-Test

When to use it

- **Comparing two groups**
- **Single dependent variable**
- **Small sample size**

When not to use it

- **Comparing more than 2 groups (use ANOVA)**
- **Non-normal Data**
- **Multiple Dependent Variables**

Analysis of Variance (ANOVA)

What is it?

ANOVA is used to compare the means of three or more groups to determine if at least one group mean is significantly different.

It partitions the total variability into between-group and within-group variability. The test statistic follows an F-distribution, which evaluates the ratio of variances.

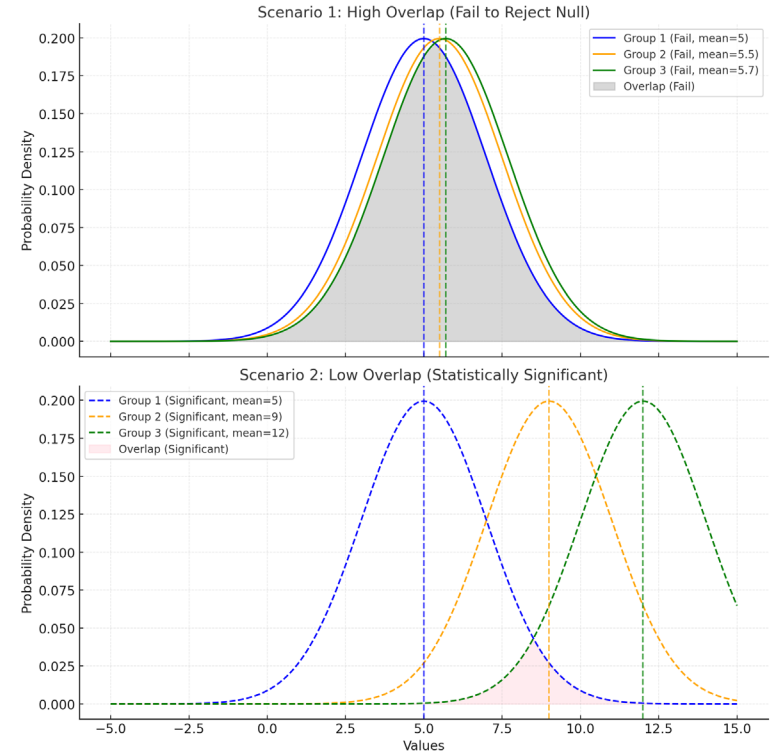
Types of ANOVA

- **One-Way ANOVA:** Compares means across a single factor with multiple groups.
- **Two-Way ANOVA:** Analyzes the effect of two factors and their interaction.
- **Repeated Measures ANOVA:** Compares means when the same subjects are measured under different conditions.

Analysis of Variance (ANOVA)

Key Calculations

1. Between-Group Variance
2. Within-Group Variance
3. Mean-Squared Between
4. Mean-Squared Within



ANOVA: Mathematical Formulation

- **Between-group variance (SSB)**

$$SSB = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

- **Mean Square Between (MSB)**

$$MSB = \frac{SSB}{k - 1}$$

- **Within-group variance (SSW)**

$$SSW = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

- **Mean Square Within (MSW)**

$$MSW = \frac{SSW}{N - k}$$

- **F-Statistic**

$$F = \frac{MSB}{MSW}$$

Where:

- k : Number of groups
- N : Overall number of observations
- n_i : Size of group i
- \bar{x}_i : Mean of group i
- \bar{x} : Overall mean of the data
- x_{ij} : Observation j in group i

ANOVA

When to use it

- **Comparing three or more groups**
- **Single dependent variable**
- **Exploring factor effects**

When not to use it

- **Fewer than 3 groups**
- **Non-normal Data**
- **Multiple Dependant Variables**



JOHNS HOPKINS

WHITING SCHOOL
of ENGINEERING

© The Johns Hopkins University 2024, All Rights Reserved.