Comparisons of Means

A direct comparison of laboratory and community EEG recordings for neurodevelopmental research

Comparison of Means in EEG Research

Example: Resting-State Alpha Power in Frontal Regions

Research Context:
Comparison of mean alpha power (8-12 Hz) between children with autism spectrum disorder (ASD) and typically developing (TD) children during resting-state EEG recordings 1.

Hypotheses:

H_0 (Null Hypothesis): The mean alpha power in frontal regions is equal in both groups ($\mu_1 = \mu_2$).

H_1 (Alternative Hypothesis): The mean alpha power in frontal regions differs between groups ($\mu_1 \neq \mu_2$).

Statistical Test:
Independent samples *t*-test (parametric test assuming normality and equal variances).

Results:

ASD Group: Mean alpha power = 4.2 μV^2

TD Group: Mean alpha power = 6.1 μV^2

p-value = 0.003

Interpretation:
The p-value (< 0.05) indicates rejection of the null hypothesis. Children with ASD show **significantly reduced alpha power** in frontal regions compared to TD children, suggesting altered neural oscillations potentially linked to differences in attention or cortical inhibition

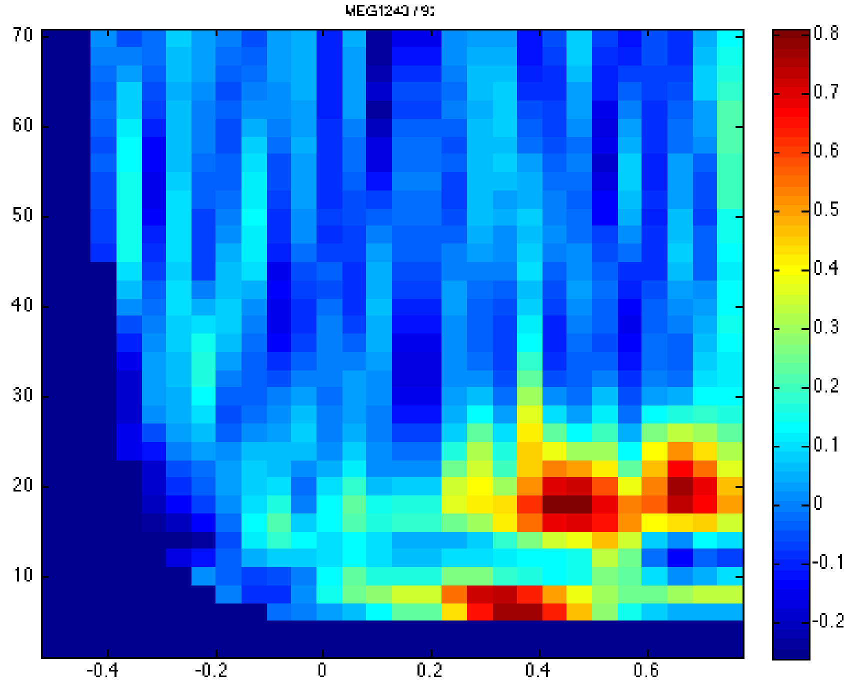
https://www.nature.com/articles/s41598-025-03569-5

Choosing the Right Statistical Test for EEG Data

Core Principle: Your choice depends on two factors:

Number of Groups: Are you comparing 2 groups (e.g., Patients vs. Controls) or more than 2 groups (e.g., Mild, Moderate, Severe ASD)?

Number of Time Points or Conditions: Are you measuring at a single time point, or repeatedly (e.g., Pre-Test vs. Post-Test, or across multiple conditions like Rest, Task1, Task2)?



Scenario 1: Two Groups, One Time Point/Condition

Research Question: **Do patients with Major Depressive Disorder (MDD) have different frontal alpha asymmetry scores than healthy controls at rest?**

Data Structure:

Between-Subjects Factor: Group (MDD vs. Control). Each participant is in only one group.

Dependent Variable (DV): A single EEG value per subject (e.g., mean frontal alpha asymmetry index).

Recommended Test: Independent Samples t-test

Example:

Comparison of Two Independent Groups

Hypotheses:

H_0: Mean alpha asymmetry is equal between MDD and control groups.

H_1: Mean alpha asymmetry differs between groups.

Statistical Test: Independent samples *t*-test.

Results:

MDD Group: Mean Asymmetry = -0.08

Control Group: Mean Asymmetry = +0.05

p-value = 0.02, Cohen's *d* = 0.7

Interpretation:

We reject the null hypothesis. The MDD group shows significantly greater relative right-frontal activity (a negative asymmetry index) compared to controls, consistent with models of withdrawal-related motivation.

Scenario 2: One Group, Two or More Time Points/Conditions

This is used when you measure the same group of people under different conditions or at different times.

Research Question: Does a mindfulness intervention alter theta power in our participants? We measure theta power during a meditation task before and after an 8-week course.

Data Structure:

Within-Subjects Factor: Time (Pre-Intervention vs. Post-Intervention). Each participant is measured twice.

Dependent Variable (DV): Theta power at each time point.

Recommended Test: Paired Samples t-test (for 2 time points) or Repeated Measures ANOVA (for 3+ time points, e.g., Pre, Mid, Post).

Example:

Comparison Within a Single Group (Repeated Measures)
Hypotheses:

H_0: Mean theta power during meditation is equal before and after the intervention.

H_1: Mean theta power during meditation changes after the intervention.

Statistical Test: Paired samples *t*-test.

Results:

Pre-Intervention: Mean Theta = 5.1 μV^2

Post-Intervention: Mean Theta = 6.8 μV^2

p-value = 0.005, Cohen's *d* = 0.9

Interpretation:
We reject the null hypothesis. Participants showed a significant increase in frontal theta power following the mindfulness training, suggesting enhanced meditative state efficiency.

Scenario 3: Two Groups, Two or More Time Points/Conditions (MIXED DESIGN)

This is one of the most common and powerful designs in experimental EEG research. It tests for an interaction effect (e.g., did one group change more over time than the other?).

Research Question: Does neurofeedback training lead to greater changes in sensorimotor rhythm (SMR) power than a sham training control?

Data Structure:

Between-Subjects Factor: Group (Neurofeedback vs. Sham).

Within-Subjects Factor: Time (Pre-Training vs. Post-Training).

Dependent Variable (DV): SMR power (12-15 Hz).

Recommended Test: Mixed-Measures ANOVA (also called Split-Plot ANOVA).

Example:

Mixed Design: Groups x Time Points

Hypotheses:

H_0: There is no interaction between Group and Time. Changes in SMR power from pre- to post-test are the same for both groups.

H_1: There is an interaction. The change in SMR power over time depends on which group participants were in.

Statistical Test: Mixed-Measures ANOVA (Group: between-subjects; Time: within-subjects).

Results:

A significant Group x Time interaction was found ($F(1, 38) = 10.5$, *p* = .002, $\eta^2 = .22$).

Simple Effects Analysis:

The Neurofeedback group showed a significant increase in SMR from Pre to Post (*p* < .001).

The Sham group showed no significant change (*p* = .45).

Interpretation:

The significant interaction confirms our experimental hypothesis. The increase in SMR power was specific to the group that received active neurofeedback training, ruling out placebo effects.

Scenario 4: Relationship Between a Continuous EEG Metric and a Continuous Behavioral Score (Correlation)

This tests if a change in an EEG measure is associated with a change in a behavioral or clinical score, without defining groups.

Research Question: Is there a relationship between the amplitude of the P300 event-related potential (ERP) and working memory performance score in a task?

Data Structure:

Variables: Two continuous variables measured for each subject.

Variable 1: P300 amplitude (μV) at electrode Pz.

Variable 2: Accuracy (%) on the working memory task.

Recommended Test: Pearson's Correlation (if data is normally distributed) or Spearman's Rank Correlation (for non-normal data or ordinal scores).

Example:

Assessing Relationships Between Continuous Variables

Hypotheses:

H_0: There is no linear relationship between P300 amplitude and working memory accuracy ($\rho = 0$).

H_1: There is a linear relationship between P300 amplitude and working memory accuracy ($\rho \neq 0$).

Statistical Test: Pearson's correlation coefficient (*r*).

Results:

Correlation coefficient (*r*) = +0.65

p-value < 0.001

$R^2 = 0.42$

Interpretation:

We reject the null hypothesis. There is a strong, statistically significant positive correlation. Participants with larger P300 amplitudes tended to have higher working memory accuracy, with the P300 amplitude explaining 42% of the variance in performance.

Scenario 5: Comparing Three or More Independent Groups

This extends Scenario 1 to more than two groups (e.g., different disorders, different dosage levels).

Research Question: Does resting-state beta power differ between healthy controls, patients with Parkinson's disease (PD), and patients with essential tremor?

Data Structure:

Between-Subjects Factor: Group (3 levels: Control, PD, Tremor). Each participant is in only one group.

Dependent Variable (DV): Mean beta power (13-30 Hz).

Recommended Test: One-Way Analysis of Variance (ANOVA)

If the ANOVA is significant (*p* < .05), you must run post-hoc tests (e.g., Tukey's HSD) to determine which specific groups differ from each other.

Example:

Comparison of More Than Two Independent Groups

Hypotheses:

H_0: Mean beta power is equal across all three groups ($\mu_1 = \mu_2 = \mu_3$).

H_1: At least one group's mean beta power is different.

Statistical Test: One-Way ANOVA.

Results:

ANOVA result: $F(2, 57) = 8.91$, *p* = 0.001

Post-hoc Tukey Tests:

Parkinson's vs. Control: *p* = .002 (PD > Control)

Tremor vs. Control: *p* = .120

Parkinson's vs. Tremor: *p* = .015 (PD > Tremor)

Interpretation:

The ANOVA was significant. Post-hoc tests revealed that the Parkinson's group had significantly elevated beta power compared to both the control group and the essential tremor group, which may reflect pathologically synchronized neural activity.

Scenario 6: Complex Within-Subjects Designs (Multiple Conditions)

This is for when you have one group but many conditions, which is very common in ERP studies.

Research Question: Does the N400 ERP component differ in amplitude across three word conditions: congruent sentences ("I drink coffee with sugar"), incongruent sentences ("I drink coffee with socks"), and meaningless sentences ("Blue theory happy orbit")?

Data Structure:

Within-Subjects Factor: Condition (3 levels: Congruent, Incongruent, Meaningless). Each participant sees all conditions.

Dependent Variable (DV): N400 mean amplitude (e.g., 300-500ms at Cz).

Recommended Test: One-Way Repeated Measures ANOVA

Followed by post-hoc tests (e.g., Bonferroni-corrected paired t-tests) if the main effect is significant.

Example:

Comparison of Multiple Within-Subject Conditions

Hypotheses:

H_0: Mean N400 amplitude is equal across all three sentence conditions.

H_1: Mean N400 amplitude differs across at least one condition.

Statistical Test: Repeated Measures ANOVA.

Results:

ANOVA result: $F(2, 38) = 45.2$, *p* $< .001$, $\eta^2 = .70$ (large effect)

Post-hoc Comparisons:

Incongruent > Congruent (*p* $< .001$)

Meaningless > Incongruent (*p* $= .003$)

Meaningless > Congruent (*p* $< .001$)

Interpretation:

We reject the null hypothesis. The N400 amplitude, a marker of semantic processing, was largest (most negative) for meaningless sentences, followed by incongruent sentences, and smallest for congruent sentences. This shows a sensitivity to both semantic violation and complete lack of meaning.

Research Question Example	Factor 1 (Between-Subjects)	Factor 2 (Within-Subjects)	Primary Statistical Test	Follow-Up Tests & Notes
Do two groups differ on one EEG measure? (e.g., ASD vs. TD on alpha power)	2 Groups	None	**Independent Samples t-test**	Report effect size (Cohen's *d*).
Do three or more groups differ on one EEG measure? (e.g., Control, PD, Tremor on beta power)	3+ Groups	None	**One-Way ANOVA**	**Post-hoc tests** (e.g., Tukey's HSD) to find which groups differ.
Does one group change across two time points/conditions? (e.g., Pre vs. Post meditation theta power)	None	2 Time Points/Conditions	**Paired Samples t-test**	Report effect size (Cohen's *d*).
Does one group change across three+ time points/conditions? (e.g., N400 across 3 sentence types)	None	3+ Time Points/Conditions	**Repeated Measures ANOVA**	**Post-hoc tests** (e.g., Bonferroni) to find which conditions differ.
Does the change over time differ between two groups? (e.g., Neurofeedback vs. Sham change in SMR power)	2 Groups	2 Time Points	**Mixed-Measures ANOVA**	**Simple Effects Analysis** to break down the interaction.
Is an EEG measure related to a behavioral score? (e.g., P300 amplitude vs. memory score)	Two Continuous Variables	N/A	**Correlation** (Pearson or Spearman)	Regression can be used for prediction.

Independent t-test

What is it?

The Independent Samples t-test (also known as the **Student's t-test** or **between-subjects t-test**) is a parametric statistical procedure used to determine whether there is a statistically significant difference between the **means of two separate, unrelated groups**.

"Independent Samples" means the participants in one group are not related to or matched with the participants in the other group. Each subject provides only one data point.

It compares the means while taking into account the variability and sample size of each group.

When to Use It

You use this test when your research design has:

One categorical independent variable (IV) with exactly two levels (e.g., Group: Patient vs. Control; Condition: Treatment A vs. Treatment B).

One continuous dependent variable (DV) that you have measured for each subject (e.g., Mean Alpha Power, P300 Amplitude, Reaction Time).

Different participants in each of the two groups.

3. Hypotheses

Null Hypothesis (H_0): The means of the two populations are equal.

$\mu_1 = \mu_2$ (There is no difference in the mean EEG measure between groups)

Alternative Hypothesis (H_1): The means of the two populations are not equal.

$\mu_1 \neq \mu_2$ (There is a difference in the mean EEG measure between groups)

This is a two-tailed test (most common). One-tailed tests ($\mu_1 > \mu_2$ or $\mu_1 < \mu_2$) are rare and require strong justification before seeing the data.

4. The Test Statistic (t-value)

The formula calculates a t-value, which is a ratio of the signal (the difference between groups) to the noise (the variability within groups).

Formula:

$t = (M_1 - M_2) / \sqrt{[(s_1^2/n_1) + (s_2^2/n_2)]}$

M_1, M_2: Means of Group 1 and Group 2

s_1^2, s_2^2: Variances of Group 1 and Group 2

n_1, n_2: Sample sizes of Group 1 and Group 2

A larger absolute t-value (further from zero) indicates a greater difference between groups relative to their internal variability.

3. Hypotheses

Null Hypothesis (H_0): The means of the two populations are equal.

$\mu_1 = \mu_2$ (There is no difference in the mean EEG measure between groups)

Alternative Hypothesis (H_1): The means of the two populations are not equal.

$\mu_1 \neq \mu_2$ (There is a difference in the mean EEG measure between groups)

This is a two-tailed test (most common). One-tailed tests ($\mu_1 > \mu_2$ or $\mu_1 < \mu_2$) are rare and require strong justification before seeing the data.

4. The Test Statistic (t-value)

The formula calculates a t-value, which is a ratio of the signal (the difference between groups) to the noise (the variability within groups).

Formula:

$t = (M_1 - M_2) / \sqrt{[(s_1^2/n_1) + (s_2^2/n_2)]}$

M_1, M_2: Means of Group 1 and Group 2

s_1^2, s_2^2: Variances of Group 1 and Group 2

n_1, n_2: Sample sizes of Group 1 and Group 2

A larger absolute t-value (further from zero) indicates a greater difference between groups relative to their internal variability.

Assumptions of the Independent Samples t-test

For the results of an independent samples t-test to be valid, the following six key assumptions must be met.

1. Two Independent Groups

2. Continuous Dependent Variable

3. Independence of Observations

4. No Significant Outliers

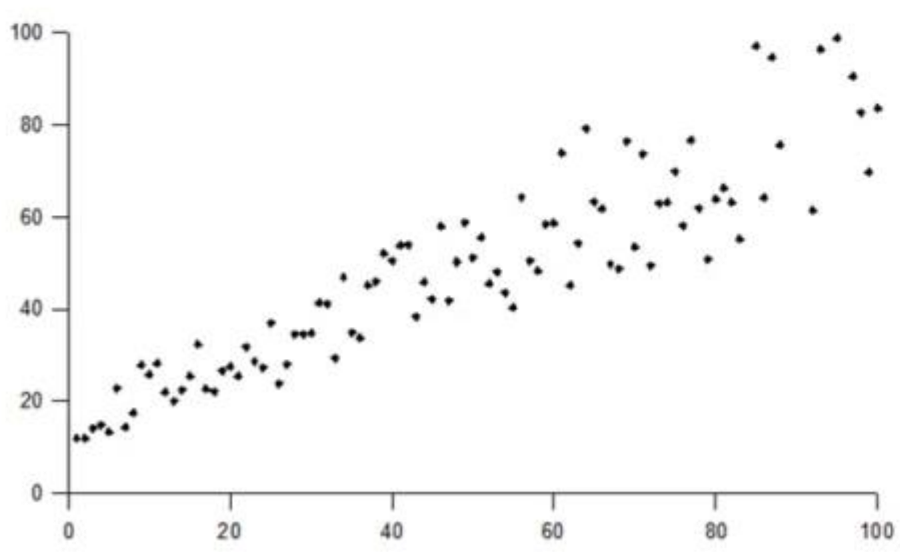
5. Normality

6. Homogeneity of Variances (Homoscedasticity)

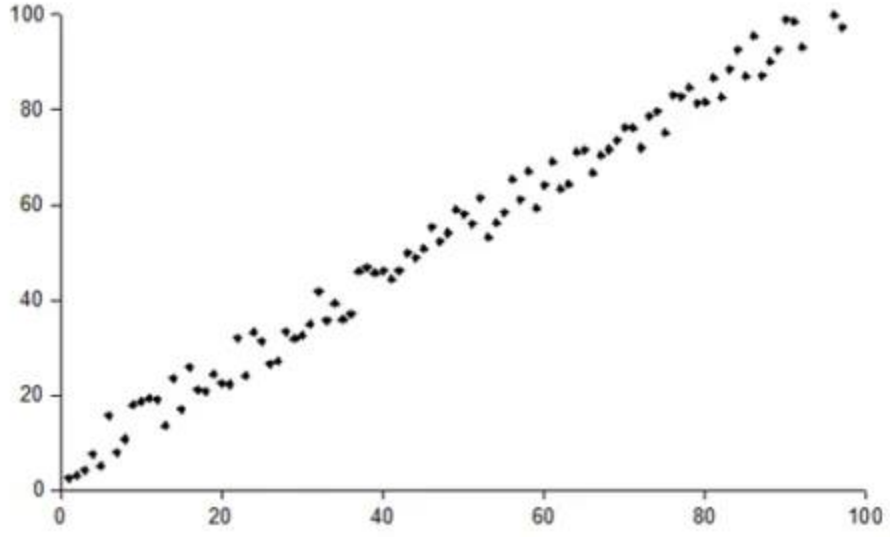
Homoscedasticity vs. Heteroscedasticity

(meaning “same variance”)

Heteroskedasticity



Homoskedasticity



Homoscedasticity vs. Heteroscedasticity

Why is it Important?

Homoscedasticity is a key assumption for many statistical tests, most notably:

Linear Regression: The reliability of the regression coefficients (their p-values and confidence intervals) depends on the assumption that the residuals (errors) have constant variance. If this assumption is violated, the standard errors can be biased, leading to unreliable significance tests.

Analysis of Variance (ANOVA): ANOVA assumes that the variances of the groups you are comparing are approximately equal. Violating this assumption can affect the Type I error rate (the chance of a false positive).

Types of T-Tests

- **One-sample T-test: Compares the mean of a single sample to a known value or hypothesized population mean**
- **Paired-sample T-test: Compares means from the same group at different times (e.g., before and after an intervention)**
- **Two-sample T-test: Compares means from two independent groups, assuming either equal or unequal variances**

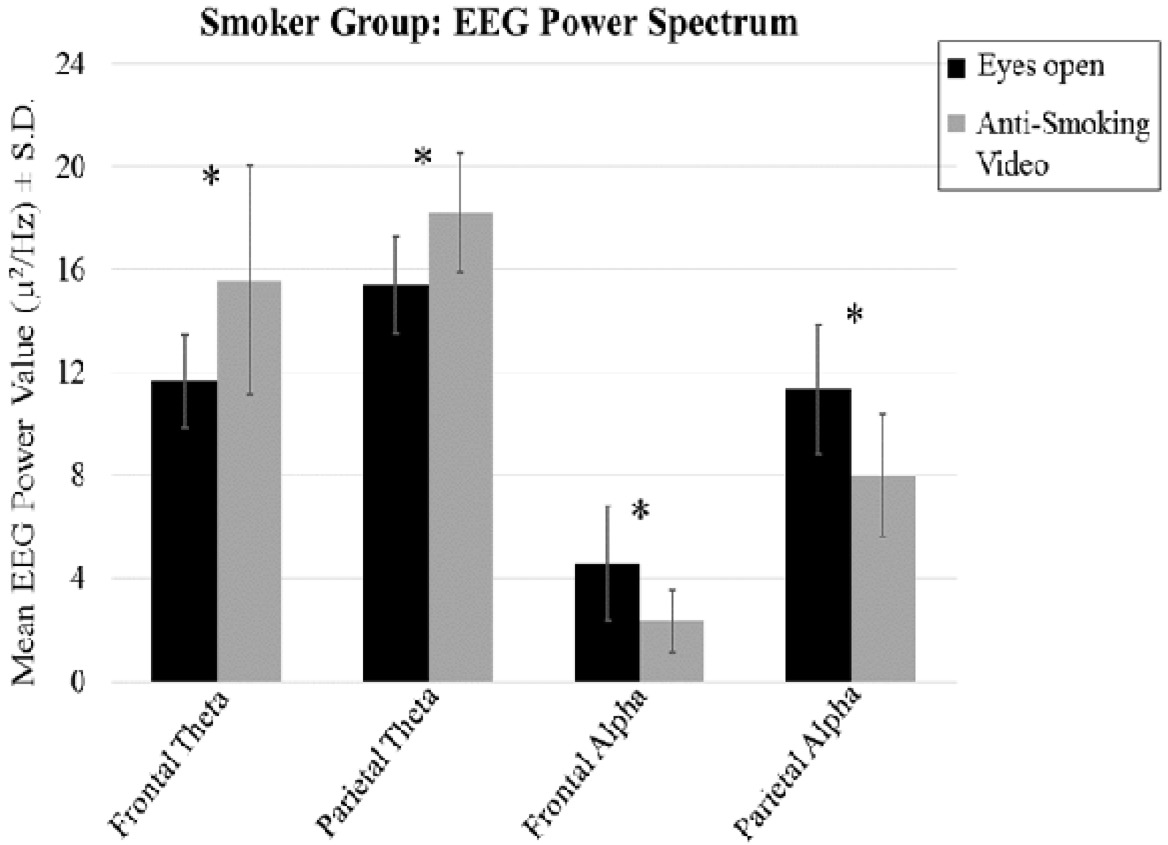
MATLAB provides built-in functions for performing T-tests:

ttest(): For one-sample and paired-sample T-tests

ttest2(): For two-sample T-tests

ranksum(): Nonparametric alternative to the two-sample T-test (Mann-Whitney U-test)

MATLAB T-TEST



Repeated measures ANOVA

What is Repeated Measures ANOVA?

Repeated Measures ANOVA is a statistical technique used to analyze the difference between the means of three or more groups where the same participants are in each group. It's the equivalent of a one-way ANOVA, but for related, not independent, groups.

The core idea is that you're measuring the same subjects repeatedly under different conditions or at different time points.

When to Use It:

You should consider a Repeated Measures ANOVA when your study design meets these criteria:

One Independent Variable (IV): You have a single categorical independent variable with three or more levels (e.g., Time: Pre-test, Mid-test, Post-test; Drug Dosage: 0mg, 50mg, 100mg; Condition: Visual, Auditory, Tactile).

Related Samples: The same participants experience all levels of the IV (you have repeated measurements).

One Continuous Dependent Variable (DV): The outcome you are measuring is continuous (e.g., Test Score, Reaction Time, Blood Pressure).

3. Hypotheses

Null Hypothesis (H_0): All population means are equal.

$\mu_1 = \mu_2 = \mu_3 \ldots = \mu_k$ (where *k* is the number of levels/conditions)

Alternative Hypothesis (H_1): At least one population mean is different from the others.

Assumptions of repeated measures ANOVA

1. Sphericity: This is the most important and unique assumption. It means the variances of the differences between all combinations of related groups are equal. It's like the homogeneity of variances assumption for a paired t-test but extended to three or more groups.

2. Continuous Dependent Variable

3. Dependent Samples

4. No Significant Outliers

5. Normality

one-Way ANOVA

What is a One-Way ANOVA?

A One-Way ANOVA (Analysis of Variance) is a statistical test used to determine if there are any statistically significant differences between the means of three or more independent (unrelated) groups defined by a single categorical independent variable.

It's an extension of the t-test, which is limited to comparing only two groups. The One-Way ANOVA is powerful because it can compare multiple groups at once, controlling the overall error rate and preventing the accumulation of false positives that could occur from performing multiple t-tests.

The Key Question it Answers

A One-Way ANOVA tests one primary hypothesis:

Main Effect: Does the single independent variable (or factor) cause a significant difference in the means of the dependent variable?

- Example: Does Diet Type (the factor with levels: Vegan, Keto, Mediterranean) have a significant effect on Weight Loss (the dependent variable)?

If the test is significant, it tells us that at least one group mean is different from the others, but it does not tell us which specific groups are different. For that, follow-up tests (known as post-hoc tests) are required.

Key Terms and Concepts:

- Factor: The single independent variable being studied (e.g., Diet Type, Drug Dosage, Marketing Campaign).
- Levels: The different categories or groups within the factor (e.g., Diet Type: Vegan, Keto, Mediterranean).
- Between-Group Variance: The variance between the different level means. This reflects the effect of the factor plus random error.
- Within-Group Variance: The variance within each level or group. This reflects random error or natural variation.
- F-ratio: The test statistic for ANOVA. It is a ratio of the Between-Group Variance to the Within-Group Variance (F = Between / Within). A larger F-ratio suggests a greater effect of the factor relative to the random noise in the data.

Assumptions of One-Way ANOVA

For the results of a One-Way ANOVA to be valid, the following assumptions must be met:

1. Independence of Observations: The data in different groups must be collected independently. This is usually achieved through random sampling or random assignment.

2. Continuous Dependent Variable: The dependent variable (what you're measuring) should be continuous (e.g., height, weight, test score, time).

3. Normality: The data within each group should be approximately normally distributed. This assumption is especially important for small sample sizes but is less critical with larger ones.

4. Homogeneity of Variances (Homoscedasticity): The variance within each of the groups should be roughly equal.

Two-Way ANOVA

What is a Two-Way ANOVA?

A Two-Way ANOVA is a statistical test used to determine the effect of two independent categorical variables (often called factors) on one continuous dependent variable.

It's an extension of the One-Way ANOVA, which only looks at one factor. The Two-Way ANOVA is powerful because it not only tells you if each factor has a significant effect but also whether there is a significant interaction effect between the two factors.

The Three Questions it Answers

A Two-Way ANOVA tests three separate hypotheses simultaneously:

Main Effect of Factor A: Does the first independent variable (e.g., Gender) cause a significant difference in the group means of the dependent variable (e.g., Test Score)?

Main Effect of Factor B: Does the second independent variable (e.g., Teaching Method) cause a significant difference in the group means?

Interaction Effect (A x B): Does the effect of one independent variable depend on the level of the other independent variable?

Example: Does the effect of Teaching Method on Test Score depend on the Gender of the student?

An interaction effect is often the most interesting part of the analysis. It can be visualized using an interaction plot.

Key Terms and Concepts:

- Factor: Another name for an independent variable (e.g., Diet, Exercise).
- Level: The different categories within a factor (e.g., Diet: Vegan, Keto, Mediterranean).
- Cell: A specific combination of the levels of the two factors (e.g., Vegan diet + High exercise).
- Main Effect: The individual effect of a single independent variable, ignoring all other independent variables.
- Interaction Effect: When the effect of one factor on the dependent variable changes depending on the level of the other factor.

Assumptions of Two-Way ANOVA

1. Dependent (Paired) Samples Across Two Factors

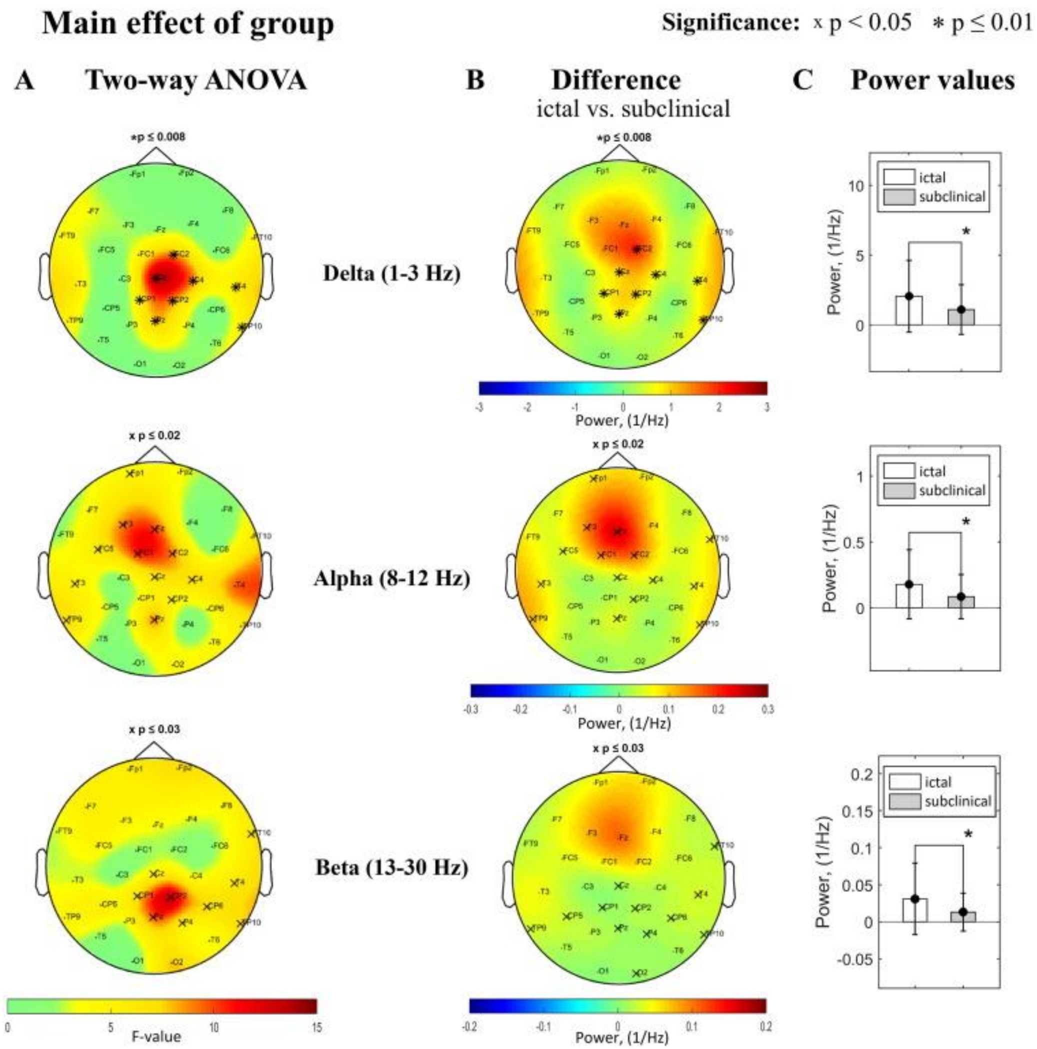
2. Continuous Dependent Variable

3. No Significant Outliers

4. Normality

5. Homogeneity of Variances

MATLAB ANOVA



Summary Table

Test	Number of Groups	Variables & Design	Purpose	Example
One-Sample T-Test	One	Compares one group's mean to a **known standard** or population mean.	Is our group different from a known value?	Is the average IQ of our sample (μ=103) different from the population mean (μ=100)?
Independent (Two-Sample) T-Test	Two	Compares means between **two independent, unrelated groups**.	Are the means of these two distinct groups different?	Do men and women have different average salaries?
Paired-Sample T-Test	Two	Compares means between **two related/matched groups**.	Did a change occur within the same subjects?	Did patients' blood pressure decrease after taking a drug? (Before vs. After)
One-Way ANOVA	Three or More	Compares means between **three or more independent groups** based on **one** factor.	Are the means of any of these multiple groups different?	Do test scores differ between students using Textbook A, B, or C?
Repeated Measures ANOVA	One or More	Compares means across **three or more related/matched time points or conditions**.	Did a change occur over multiple measurements?	Did patient anxiety change over time (Baseline, 3 months, 6 months)?
Two-Way ANOVA	Multiple Groups	Examines the effect of **two independent factors** on a dependent variable.	1. What is the effect of each factor? 2. Is there an **interaction** between them?	Do both Gender (Factor A) and Diet (Factor B) affect weight loss, and does Diet work differently for men and women?