A one-way ANOVA (“analysis of variance”) is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups.

Example: One-Way ANOVA in Python A researcher recruits 30 students to participate in a study.

The students are randomly assigned to use one of three studying techniques for the next three weeks to prepare for an exam. At the end of the three weeks, all of the students take the same test.

Use the following steps to perform a one-way ANOVA to determine if the average scores are the same across all three groups.

Example1:

Step 1: Enter the data.

First, we’ll enter the exam scores for each group into three separate arrays:

enter exam scores for each group

group1 = [85, 86, 88, 75, 78, 94, 98, 79, 71, 80]

group2 = [91, 92, 93, 85, 87, 84, 82, 88, 95, 96]

group3 = [79, 78, 88, 94, 92, 85, 83, 85, 82, 81]

```
group1 = [85, 86, 88, 75, 78, 94, 98, 79, 71, 80]

group2 = [91, 92, 93, 85, 87, 84, 82, 88, 95, 96]

group3 = [79, 78, 88, 94, 92, 85, 83, 85, 82, 81]
```

Step 2: Perform the one-way ANOVA.

Next, we'll use the f_oneway() function from the SciPy library to perform the one-way ANOVA:

```
from scipy.stats import f_oneway

#perform one-way ANOVA
f_oneway(group1, group2, group3)

#(statistic=2.3575, pvalue=0.1138)
```

```
F_onewayResult(statistic=2.3575322551335636, pvalue=0.11384795345837218)
```

Step 3: Interpret the results.

A one-way ANOVA uses the following null and alternative hypotheses:

H0 (null hypothesis): μ1 = μ2 = μ3 = … = μk (all the population means are equal)

H1 (null hypothesis): at least one population mean is different from the rest

The F test statistic is 2.3575 and the corresponding p-value is 0.1138.

Since the p-value is not less than .05, we fail to reject the null hypothesis.

This means we do not have sufficient evidence to say that there is a difference in exam scores among the three studying techniques.

Example 2: The average score on a test is 80 with a standard deviation of 10. With a new teaching curriculum introduced it is believed that this score will change. On random testing, the score of 38 students, the mean was found to be 88. With a 0.05 significance level, is there any evidence to support this claim?

Solution: This is an example of two-tail hypothesis testing. The z test will be used.

$H_0: \mu = 80$, $H_1: \mu \neq 80$

$\bar{x} = 88$, $\mu = 80$, $n = 36$, $\sigma = 10$.

$\alpha = 0.05 / 2 = 0.025$

The critical value using the normal distribution table is 1.96

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{88 - 80}{\frac{10}{\sqrt{36}}} = 4.8$$

As 4.8 > 1.96, the null hypothesis is rejected.

Answer: There is a difference in the scores after the new curriculum was introduced.

Example 3: The average score of a class is 90. However, a teacher believes that the average score might be lower. The scores of 6 students were randomly measured. The mean was 82 with a standard deviation of 18. With a 0.05 significance level use hypothesis testing to check if this claim is true.

Solution: The t test will be used.

$H_0: \mu = 90, H_1: \mu < 90$

$\bar{x} = 110, \mu = 90, n = 6, s = 18$

The critical value from the t table is -2.015

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{82 - 90}{\frac{18}{\sqrt{6}}}$$

$t = -1.088$

As -1.088 > -2.015, we fail to reject the null hypothesis.

Answer: There is not enough evidence to support the claim.

Double-click (or enter) to edit

One-way ANOVA | When and How to Use It (With Examples)

ANOVA, which stands for Analysis of Variance, is a statistical test used to analyze the difference between the means of more than two groups.

A one-way ANOVA uses one independent variable, while a two-way ANOVA uses two independent variables.

One-way ANOVA example

As a crop researcher, we want to test the effect of three different fertilizer mixtures on crop yield.

we can use a one-way ANOVA to find out if there is a difference in crop yields between the three groups.

Table of contents

When to use a one-way ANOVA

Use a one-way ANOVA when we have collected data about one categorical independent variable and one quantitative dependent variable.

The independent variable should have at least three levels (i.e. at least three different groups or categories).

ANOVA tells us if the dependent variable changes according to the level of the independent variable. For example:

our independent variable is social media use, and we assign groups to low, medium, and high levels of social media use to find out if there is a difference in hours of sleep per night.

our independent variable is brand of soda, and we collect data on Coke, Pepsi, Sprite, and Fanta to find out if there is a difference in the price per 100ml.

The null hypothesis (H0) of ANOVA is that there is no difference among group means.

The alternate hypothesis (Ha) is that at least one group differs significantly from the overall mean of the dependent variable.

If we only want to compare two groups, use a t-test instead.

How does an ANOVA test work?

ANOVA determines whether the groups created by the levels of the independent variable are statistically different by calculating whether the means of the treatment levels are different from the overall mean of the dependent variable.

If any of the group means is significantly different from the overall mean, then the null hypothesis is rejected.

ANOVA uses the F-test for statistical significance. This allows for comparison of multiple means at once, because the error is calculated for the whole set of comparisons rather than for each individual two-way comparison (which would happen with a t-test).

The F-test compares the variance in each group mean from the overall group variance.

If the variance within groups is smaller than the variance between groups, the F-test will find a higher F-value, and therefore a higher likelihood that the difference observed is real and not due to chance.