A t-test is a statistical test used to compare the means of two groups to determine if they are statistically different from each other. It’s useful when dealing with small sample sizes and is based on the t-distribution.

The *t-test* estimates the true difference between two group means using the ratio of the difference in group means over the pooled standard error of both groups. A larger t-value shows that the difference between group means is greater than the pooled standard error, indicating a more significant difference between the groups. You can compare your calculated *t-value* against the values in a critical value chart (e.g., Student’s *t* table) to determine whether your *t-value* is greater than what would be expected by chance. If so, the null hypothesis can be rejected and the alternate hypothesis is true.

A t-test can only be used when comparing the means of **two** groups (a.k.a. pairwise comparison). If you want to compare more than two groups, or if you want to do multiple pairwise comparisons, use an ANOVA test or a post-hoc test.

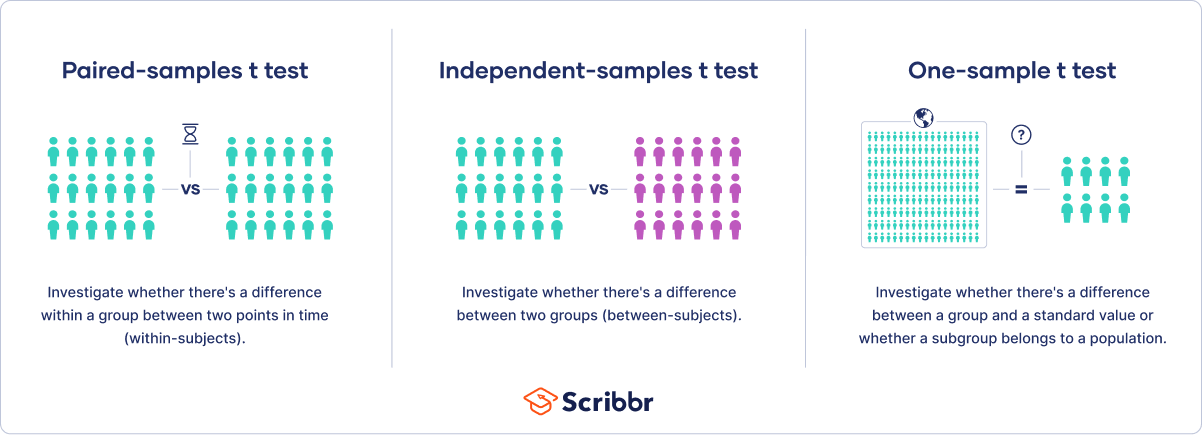
The t-test is a parametric test of difference, meaning that it makes the same assumptions about your data as other parametric tests. The t-test assumes your data:

* Are independent
* Are (approximately) normally distributed
* Have a similar amount of variance within each group being compared (a.k.a. homogeneity of variance)

If your data don’t fit these assumptions, you can try a nonparametric alternative to the t-test, such as the Wilcoxon Signed-Rank test for data with unequal variances.

**What type of *t-test* should be used?**

When choosing a *t-test*, consider two things: whether the groups being compared come from a single population or two different populations, and whether you want to test the difference in a specific direction.



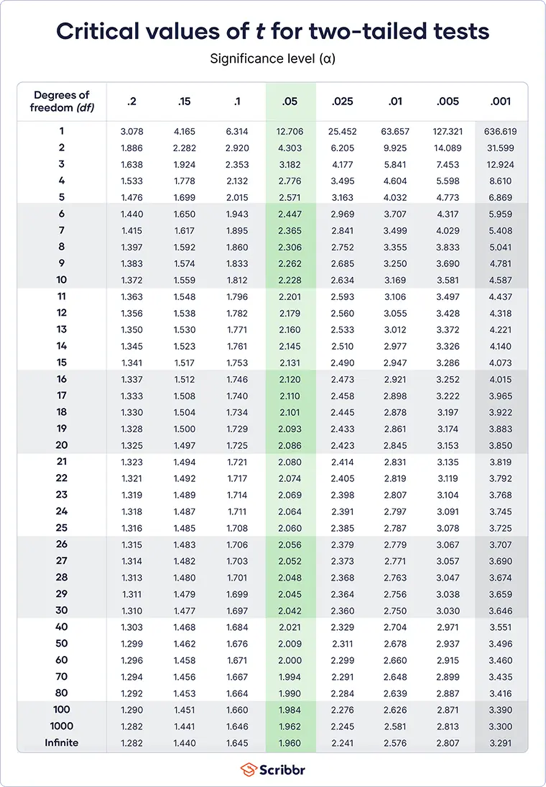
**Types of t-tests:**

1. **Independent t-test**: Compares means from two different populations (e.g., amount of leaf growth in two species of plants).
2. **Paired t-test**: Compares means from the same population at different times (e.g., weight loss measurements before and after a diet).
3. **One-sample t-test**: Compares the mean of a single group against a known value or population mean.

**One-tailed or two-tailed *t-test*?**

* If you only care whether the two populations are different from one another, perform a **two-tailed *t-test***.
* If you want to know whether one population mean is greater than or less than the other, perform a **one-tailed *t-test*.**
* We’ll be covering the two-tailed t-test.

**Steps to perform a t-test:**

1. **Formulate Hypotheses (THE MOST IMPORTANT PART)**:
   * **Null Hypothesis (H0)**: Assumes no difference between the group means.
   * **Alternative Hypothesis (H1)**: Assumes there is a difference.
2. **Collect Data**: Gather your data for the two groups you wish to compare.
3. **Calculate the t-statistic**:
   * For an independent t-test, the formula is:
   * Where:
     + and are the sample means.
     + and ​ are the sample variances (standard error/deviation).
     + ​ and are the sample sizes.
4. **Determine Degrees of Freedom**
   * For an independent t-test:
5. **Find the Critical t-value**: Use a t-distribution table to find the critical value based on your chosen significance level (usually 0.05) and degrees of freedom.
6. **Decide**:
   * If the absolute value of the calculated t-statistic is greater than the critical t-value, reject the null hypothesis.
7. **Calculate the p-value**: Alternatively, you can calculate the p-value and compare it to your significance level.
8. **Report Results**: Include the t-statistic, degrees of freedom, p-value, and a conclusion regarding the hypotheses.

**Example of an Independent t-test:**

1. **Hypotheses**:
   * H0:
   * H1:
2. **Data**: Assume Group A has scores [80, 85, 90] and Group B has scores [75, 70, 65].
3. **Calculate means and variances**:
   * ,
   * , ​
   * ​,
4. **Calculate t**:
5. **Degrees of Freedom**:
6. **Find critical t-value** for at α=0.05 (two-tailed) ≈ 2.776.
7. **Conclusion**: Since 1.71 < 2.776 - > fail to reject H0; no significant difference in means.

To do t-tests in python:

* Install scipy and numpy
* From scipy import stats
* Prepare two data sets into to separate groups (i.e. np.array)
* Run t-test:

[t\_stat, p\_value = stats.ttest\_ind(group\_a, group\_b)]

* For a paired t-test:

[t\_stat, p\_value = stats.ttest\_rel(before, after)]

* For a one-sample t-test:

[t\_stat, p\_value = stats.ttest\_1samp(sample\_data, population\_mean)]

* If p\_value is less than a set standard (i.e. alpha), reject the null. Otherwise fail to reject the null