

# The Complete Guide: Hypothesis Testing in R

Authored by  
**Mohammed looti**

November 4, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *The Complete Guide: Hypothesis Testing in R*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=9633>

A [Hypothesis Test](#) is the cornerstone of quantitative analysis, providing a structured, formal statistical procedure to evaluate claims about population parameters. The core goal is to determine, based on sample evidence, whether we possess sufficient reason to reject a predefined assumption, known as the null hypothesis. This rigorous approach is absolutely fundamental to [statistical inference](#) and drives evidence-based decision-making across all fields of data science.

The [R programming language](#) is the industry standard for statistical computation, offering powerful and intuitive functions for executing these tests efficiently. This comprehensive tutorial focuses specifically on the family of [Student's t-tests](#)--essential tools for comparing means. We will walk through the implementation of three critical variants of the t-test using R:

One-Sample T-Test (Comparing a sample mean to a fixed value)

Two-Sample T-Test (Comparing means of two independent groups)

[Paired Samples T-Test](#) (Comparing means of dependent, related groups)

## The Universal T-Test Function: `t.test()`

The efficiency of statistical testing in R stems largely from the versatile `t.test()` function. This single function handles all three major types of t-tests--one-sample, independent two-sample, and paired--simply by adjusting its parameters. Mastering these arguments is crucial for accurate statistical analysis, whether you are comparing a single sample mean against a target value or evaluating the difference between two sample means.

### # General syntax for the `t.test` function in R

```
t.test(x, y = NULL,  
alternative = c("two.sided", "less", "greater"),  
mu = 0, paired = FALSE, var.equal = FALSE,  
conf.level = 0.95, ...)
```

The behavior and output of the test are primarily governed by the following core parameters:

**x, y:** These are the numerical data vectors representing the sample(s). The `y` argument is omitted when running a one-sample test.

**alternative:** This defines the directionality of the test based on the [alternative hypothesis](#). Options include `"two.sided"` (the mean is not equal to the hypothesized value), `"less"`, or `"greater"`.

**mu:** Represents the hypothesized true value of the mean under the null hypothesis. It is primarily used when conducting a one-sample test, where the default value is zero.

**paired:** Setting this logical parameter to `TRUE` tells R that the observations in vector `x` are dependent upon, or matched with, the observations in vector `y`, thereby executing a [paired t-test](#)

instead of an independent samples test.

**var.equal:** A logical indicator (`TRUE` or `FALSE`) that determines whether the test assumes equal population [variance](#) between the two samples. If set to `FALSE` (the default), R employs [Welch's t-test](#), which is generally safer when variances are unknown or unequal.

**conf.level:** Specifies the desired [confidence level](#) for the interval estimate of the population mean(s). The standard default is 0.95 (or 95%).

Now that we understand the mechanics of the `t.test()` function, let us proceed to practical examples demonstrating how to apply this powerful function for different statistical testing scenarios.

## Example 1: Performing a One-Sample T-Test in R

The [one-sample t-test](#) is the appropriate tool when researchers want to evaluate whether the mean of a single population differs significantly from a predefined constant value, often denoted as  $\mu_0$ . This test is fundamental for validating assumptions about a population based on limited sample data.

We will use a classic scenario involving ecological data. Suppose wildlife biologists hypothesize that the average weight of a particular turtle species is 310 pounds. To investigate this claim, they collect a simple random sample of 13 turtles and record their weights. We must test the [null hypothesis](#) ( $H_0$ :  $\mu = 310$ ) against the two-sided alternative hypothesis ( $H_a$ :  $\mu \neq 310$ ).

The collected sample weights (in pounds) are: **300, 315, 320, 311, 314, 309, 300, 308, 305, 303, 305, 301, 303**. The R code below demonstrates the execution of the one-sample t-test. Notice how the `mu` argument is essential here, as it explicitly defines the hypothesized population mean against which the sample is compared.

```
#define vector of turtle weights
```

```
turtle_weights <- c(300, 315, 320, 311, 314, 309, 300, 308, 305, 303, 305, 301, 303)
```

```
#perform one sample t-test against mu = 310
```

```
t.test(x = turtle_weights, mu = 310)
```

```
One Sample t-test
```

```
data: turtle_weights
```

```
t = -1.5848, df = 12, p-value = 0.139
```

```
alternative hypothesis: true mean is not equal to 310
```

```
95 percent confidence interval:
```

303.4236 311.0379

sample estimates:

mean of x

307.2308

To interpret these results, we focus on the key statistical output provided by R:

T-test statistic: **-1.5848**. This measures how many standard errors the sample mean (307.2308) is away from the hypothesized mean (310).

**Degrees of Freedom (df): 12** (Calculated as  $n-1$ , where  $n$  is the sample size).

**P-value: 0.139**. This is the probability of observing our data (or more extreme data) if the null hypothesis were true.

95% **Confidence Interval** for the true mean: . Note that the hypothesized value of 310 falls within this range.

Sample Mean: **307.2308**.

The crucial step is drawing a conclusion based on the [p-value](#). Since 0.139 is significantly greater than the conventional significance level ( $\alpha = 0.05$ ), we must fail to reject the [null hypothesis](#). Therefore, we conclude that there is insufficient statistical evidence from this sample to claim that the true average weight of the turtle species differs from 310 pounds.

## Example 2: Independent Two-Sample T-Test in R

The **Two-Sample T-Test**, often referred to as the independent samples t-test, is employed when the objective is to assess whether a significant statistical difference exists between the means of two distinct and unrelated populations. This test is invaluable in experimental design, such as comparing the effectiveness of two different treatments or analyzing performance metrics across two separate demographic groups.

Continuing our ecological study, let us now compare the average weights of two entirely different turtle species, Species A and Species B. The [null hypothesis](#) for this test is that the population mean weight of Species A ( $\mu_1$ ) is equal to the population mean weight of Species B ( $\mu_2$ ), meaning  $H_0: \mu_1 = \mu_2$ . The alternative hypothesis is that they are not equal.

We collected separate, independent random samples for each species. Since the sample sizes are equal but the underlying population variances are unknown, the most robust approach is to run the test without assuming equal variance. R handles this automatically by executing the [Welch Two Sample t-test](#) by default (i.e., setting `var.equal = FALSE`).

The data vectors are defined as follows:

**Sample 1 (Species A):** 300, 315, 320, 311, 314, 309, 300, 308, 305, 303, 305, 301, 303

**Sample 2 (Species B):** 335, 329, 322, 321, 324, 319, 304, 308, 305, 311, 307, 300, 305

The code below passes both vectors, `x` and `y`, to the `t.test()` function:

```
#define vector of turtle weights for each sample
sample1 <- c(300, 315, 320, 311, 314, 309, 300, 308, 305, 303, 305, 301, 303)
sample2 <- c(335, 329, 322, 321, 324, 319, 304, 308, 305, 311, 307, 300, 305)

#perform two sample t-test (Welch's test by default)
t.test(x = sample1, y = sample2)
```

Welch Two Sample t-test

```
data: sample1 and sample2
t = -2.1009, df = 19.112, p-value = 0.04914
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-14.73862953 -0.03060124
sample estimates:
mean of x mean of y
307.2308 314.6154
```

The output reveals the following key findings:

T-test statistic: **-2.1009**.

**Degrees of Freedom: 19.112.** Note that the fractional degrees of freedom are characteristic of Welch's method, which adjusts for unequal [variance](#).

**P-value: 0.04914.**

95% [Confidence Interval](#) for true mean difference ( $\mu_1 - \mu_2$ ): . Since this interval does not contain zero, it strongly suggests a difference exists.

Sample Means: Species A mean is **307.2308**; Species B mean is **314.6154**.

The calculated [p-value](#) (0.04914) is marginally less than the conventional significance threshold of  $\alpha = 0.05$ . Based on this result, we successfully reject the [null hypothesis](#). This statistical conclusion provides evidence that the true mean weights of the two turtle species are indeed significantly different from one another.

### Example 3: Analyzing Dependent Data with the Paired T-Test

The **Paired Samples T-Test** is a specialized statistical procedure mandatory for situations where data points are naturally linked or dependent. This often arises in longitudinal studies, such as "pre-test/post-test" designs, or when comparing measurements taken from matched pairs (e.g., twins or spouses). Unlike the independent two-sample test, the paired test focuses on the mean difference between the pairs, eliminating variability between subjects and increasing statistical power.

Let us examine the effectiveness of a new athletic training program. We are interested in whether this program increases the maximum vertical jump height (measured in inches) of basketball players. We recruit 12 players and record their jump height before the program implementation (Baseline) and again after one month of intensive training (Post-intervention).

The collected measurements are intrinsically linked, as the 'After' measurement corresponds directly to the same individual who provided the 'Before' measurement. To correctly analyze this dependent structure, we must utilize the `paired = TRUE` parameter within the `t.test()` function. Our **null hypothesis** is that the mean difference in jump height is zero.

The raw data for the intervention study is as follows:

**Before Program (Inches):** 22, 24, 20, 19, 19, 20, 22, 25, 24, 23, 22, 21

**After Program (Inches):** 23, 25, 20, 24, 18, 22, 23, 28, 24, 25, 24, 20

The execution of the paired t-test in R is straightforward, as shown below:

```
#define before and after max jump heights
before <- c(22, 24, 20, 19, 19, 20, 22, 25, 24, 23, 22, 21)
after <- c(23, 25, 20, 24, 18, 22, 23, 28, 24, 25, 24, 20)
```

```
#perform paired samples t-test
t.test(x = before, y = after, paired = TRUE)
```

Paired t-test

data: before and after

t = -2.5289, df = 11, p-value = 0.02803

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.3379151 -0.1620849

sample estimates:

mean of the differences

-1.25

Upon reviewing the output, we extract the critical statistical metrics:

T-test statistic: **-2.5289**.

**Degrees of Freedom**: 11 (Calculated as  $n-1$ , where  $n$  is the number of pairs).

**P-value**: **0.02803**.

95% **Confidence Interval** for true mean difference: . Since this entire interval is negative and excludes zero, it reinforces the finding of a significant increase in jump height.

Mean of the differences (Before - After): **-1.25**. This shows that, on average, the jump height increased by 1.25 inches after the training program.

Because the **p-value** (0.02803) is less than the significance level ( $\alpha = 0.05$ ), we reject the **null hypothesis**. This statistically significant result confirms that the training program successfully led to a measurable improvement in the maximum vertical jump height among the basketball players.

## Summary and Next Steps in Hypothesis Testing

This guide has provided a practical, code-focused approach to executing the three most common variations of the t-test using R's powerful `t.test()` function. We have demonstrated how to handle single-sample comparisons, independent two-sample comparisons (using **Welch's t-test** by default), and crucial paired-sample comparisons, emphasizing the interpretation of the **t-statistic**, **degrees of freedom**, and the final decision based on the **confidence interval** and p-value.

As you continue your journey in statistical analysis, remember that while R automates the calculation, the responsibility lies with the researcher to correctly identify the type of test required (e.g., distinguishing between independent and paired samples) and to accurately state the **null hypothesis** and **alternative hypothesis**.

## Additional Resources for Statistical Testing

To deepen your understanding or to quickly verify the results of your R analysis against standard tools, we recommend exploring further documentation and online statistical calculators designed to automatically perform various t-tests: