

How to Easily Perform Hypothesis Testing in R

Authored by
stats writer

December 4, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform Hypothesis Testing in R*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=105352>

Executing a hypothesis test in the R programming environment requires careful setup. Initially, you must define the specific type of test, identify the dataset, and clearly formulate both the null and alternative hypotheses. Once these prerequisites are met, R provides powerful built-in functions to compute the essential elements of the test, namely the test statistic and the corresponding p-value. The final step involves drawing a statistical conclusion by comparing the calculated p-value against the predefined significance level (often denoted as alpha).

A statistical hypothesis test is a structured procedure used to determine whether there is enough evidence in a sample of data to reject a specific hypothesis about a population parameter.

In this comprehensive tutorial, we will focus specifically on conducting different variations of the t-test in R, a fundamental statistical tool for comparing means. We will cover three essential applications:

The One Sample t-test, used for comparing a sample mean to a known value.

The Two Sample t-test (Independent samples), used for comparing the means of two distinct groups.

The Paired Samples t-test, used when measurements are taken twice on the same subjects or matched pairs.

All these variations are implemented efficiently using the built-in R function, **t.test()**. Understanding its syntax is key to proper execution.

#General syntax for the t.test() function

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

The primary arguments for the **t.test()** function are defined as follows:

x, y: These represent the numerical vectors containing the data from the first sample (**x**) and, optionally, the second sample (**y**) if performing a two-sample test.

alternative: Specifies the nature of the alternative hypothesis, determining if the test is two-sided (default), lower-tailed ("less"), or upper-tailed ("greater").

mu: The hypothesized population mean under the null hypothesis. This is primarily used for one-sample tests.

paired: A logical value (TRUE/FALSE) indicating whether the observations are paired. Setting this to **TRUE** performs a paired samples t-test.

var.equal: A logical value (TRUE/FALSE) indicating whether to assume the population variance is

equal between the samples. By default, R performs Welch's t-test (assuming unequal variance).

conf.level: The confidence level (e.g., 0.95 for a 95% confidence interval) to use for the resulting interval estimate.

The subsequent examples illustrate practical scenarios for employing this powerful function in statistical analysis.

Example 1: Analyzing a Single Sample Mean with the One Sample T-Test

The purpose of a One Sample t-test is to statistically determine if the mean of a population differs significantly from a predefined, hypothesized value. This test is foundational when assessing a single group against a known benchmark or target.

Consider a scenario where researchers are studying a specific species of turtle. We hypothesize that the population mean weight (μ) is 310 pounds. To test this hypothesis ($H_0: \mu = 310$ vs. $H_a: \mu \neq 310$), we collect a simple random sample of 13 turtles. The recorded weights for this sample are provided below:

Weights (in pounds): 300, 315, 320, 311, 314, 309, 300, 308, 305, 303, 305, 301, 303

We utilize the **t.test()** function in R, specifying our data vector (x) and the hypothesized mean (mu), to execute the test. The default setting performs a two-sided test, assuming we are interested in deviations both above and below 310 pounds.

#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 hypothesized mean of 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

Interpreting the output generated by R reveals several critical components of the test results:

t-test statistic: **-1.5848**. This measures how many standard errors the sample mean is from the hypothesized mean.

degrees of freedom (df): **12**. Calculated as $N - 1$ ($13 - 1 = 12$).

p-value: **0.139**. This is the probability of observing the data, or data more extreme, if the null hypothesis were true.

95% confidence interval for true mean: . We are 95% confident that the true population mean lies within this range.

Sample mean of turtle weights: **307.2308**.

Our critical conclusion relies on the p-value. Since the calculated p-value of 0.139 is greater than the standard significance level ($\alpha = 0.05$), we lack sufficient evidence to reject the null hypothesis. Therefore, we conclude that the mean weight of this turtle species is not statistically different from 310 pounds.

Example 2: Comparing Two Independent Means with the Two Sample T-Test

The Two Sample t-test, also known as the independent samples t-test, is employed when the goal is to assess whether the mean values of two distinct, unrelated populations are statistically equivalent. This is critical for comparative studies where data points in one group do not influence the other.

Suppose we are now comparing the weights of two separate turtle species (Species A and Species B). Our objective is to determine if there is a statistically significant difference in their average weights. We establish the null hypothesis as $H_0: \mu_A = \mu_B$ (no difference in means) and the alternative as $H_1: \mu_A \neq \mu_B$. We gather samples from both species:

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

To run this comparison in R, we pass both sample vectors (x and y) to the **t.test()** function. Since we did not specify the `var.equal = TRUE` argument, R automatically executes Welch's t-test, which accounts for the possibility of unequal variances between the two samples.

#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 t-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 results of the analysis provide the necessary statistics to make a decision:

t-test statistic: **-2.1009**. This indicates the observed difference between the sample means.

degrees of freedom (df): **19.112**. Note that for Welch's t-test, the degrees of freedom are often non-integer.

p-value: **0.04914**.

95% confidence interval for true mean difference: . Since this interval does not contain zero, it suggests a significant difference.

Mean of Sample 1 weights: **307.2308**.

Mean of Sample 2 weights: **314.6154**.

We observe that the p-value (0.04914) is marginally less than the conventional alpha level of 0.05. Consequently, we must reject the null hypothesis.

This statistically significant finding provides compelling evidence to conclude that the mean weight between the two turtle species is, in fact, unequal.

Example 3: Assessing Change Over Time with the Paired Samples T-Test

The Paired Samples t-test is the appropriate statistical tool when measurements are taken from the same subjects or matched units under two different conditions--a typical design for "before-and-after" studies. This test focuses on the mean difference between the pairs, not the means of the samples themselves.

We want to evaluate the effectiveness of a specialized training regimen designed to improve the maximum vertical jump height (measured in inches) of basketball players. We recruit 12 college athletes and measure their initial jump heights. After they complete the one-month training program, we measure their heights again.

The core hypothesis is $H_0: \mu_{\text{difference}} = 0$ (the training had no effect) versus $H_a: \mu_{\text{difference}} \neq 0$?

0 (the training changed the mean jump height). The collected data reflecting jump heights before and after the program are:

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

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

To properly conduct this analysis in R, we must explicitly set the `paired` argument to **TRUE** within the `t.test()` function, indicating that the observations in the `before` vector correspond directly to those in the `after` vector.

#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

The output focuses on the difference between the two samples:

t-test statistic: **-2.5289**.

degrees of freedom (df): **11**. Calculated as $N - 1$ (12 pairs - 1 = 11).

p-value: **0.02803**.

95% confidence interval for true mean difference: . Since both bounds are negative, the "after" mean is significantly higher than the "before" mean.

Mean difference between before and after: **-1.25**. The negative sign suggests the average jump increased by 1.25 inches after the program.

Given that the p-value (0.02803) is less than the alpha level (0.05), we successfully reject the null hypothesis.

This result confirms that we have sufficient statistical evidence to assert that the training program resulted in a significant change--specifically, an increase--in the mean vertical jump height of the basketball players.

Conclusion and Further Resources

The **t.test()** function in R provides a streamlined and reliable method for performing the most common types of t-tests--one sample, two sample, and paired samples--by simply adjusting the arguments `mu`, `y`, and `paired`. Mastering the interpretation of the resulting t-statistic and p-value is essential for rigorous statistical reporting.

While R offers robust command-line analysis, if you require quick verification or alternative tools for calculation, you may find the following resources useful.

Use the following online calculators to automatically perform various t-tests: