

How to Easily Interpret T-Test Results in R

Authored by
stats writer

December 5, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Interpret T-Test Results in R*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=105639>

When approaching the interpretation of t-test results within the R statistical environment, analysts must focus on several core metrics, primarily the p-value and the t-statistic. The p-value represents the probability of observing data as extreme as, or more extreme than, the actual observations, assuming that the null hypothesis is true. This value is paramount in determining statistical significance.

If this calculated p-value falls below the predetermined significance level (alpha, conventionally set at 0.05), we possess sufficient statistical evidence to reject the null hypothesis. The t-statistic, conversely, serves as a measure of the magnitude of the observed difference relative to the variation in the sample data; it quantifies the distance between the sample means in terms of standard error units. A larger absolute value for the t-statistic implies that the means of the two groups are substantially further apart, providing stronger evidence against the null hypothesis.

A t-test is a fundamental inferential statistical technique used to determine whether there is a significant difference between the means of two groups, which are presumed to be drawn from normally distributed populations with unknown variances.

This comprehensive guide offers a step-by-step methodology for executing and thoroughly interpreting the numerical output of a two-sample t-test when utilizing the statistical capabilities provided by R. Mastering this interpretation is crucial for drawing valid, data-driven conclusions in comparative studies.

Understanding the Foundational Concepts

Before diving into the coding, it is essential to establish the formal hypotheses being tested. In a two-sample comparison, the primary goal is often to determine if the population mean of Group 1 (μ_1) is different from the population mean of Group 2 (μ_2). This forms the basis of the competing hypotheses.

The null hypothesis, denoted as H_0 , posits that there is no difference between the population means ($\mu_1 = \mu_2$). Conversely, the alternative hypothesis, H_A , which we seek to prove, states that a difference does exist ($\mu_1 \neq \mu_2$). The interpretation of the p-value will directly guide our decision regarding the rejection or retention of H_0 .

Furthermore, the statistical power of the test relies heavily on factors such as sample size, the magnitude of the effect, and the variability within the data. Ensuring that the data meets the necessary assumptions--such as independence of observations, approximate normality, and, depending on the test type, equality of variances--is vital for the validity of the final statistical inference. Failure to meet these assumptions might necessitate the use of non-parametric alternatives.

Step 1: Defining and Preparing the Data Vectors in R

To illustrate the process, let us construct a hypothetical scenario. Imagine we are agricultural researchers investigating whether two distinct species of plants (Species A and Species B) exhibit the same average height. To conduct this investigation rigorously, we gather a representative simple random sample, collecting height measurements from 12 individual plants for each species. This careful collection process ensures that our subsequent statistical analysis is based on unbiased data.

These measured heights must be input into R as separate numeric vectors. In the code block below, we define ``group1`` (representing Species A) and ``group2`` (representing Species B). Data preparation in this manner is the crucial first step, providing the necessary input structure for the ``t.test()`` function, which is designed to handle comparative analyses between two independent sample groups.

The data vectors below contain the actual height measurements in a specified unit (e.g., centimeters). Note the slight visible difference in means, which the t-test will formally assess for statistical significance, determining if this observed difference is likely due to genuine population differences or merely random sampling variation.

```
#create vector of plant heights from group 1 (Species A)  
group1 <- c(8, 8, 9, 9, 9, 11, 12, 13, 13, 14, 15, 19)
```

```
#create vector of plant heights from group 2 (Species B)  
group2 <- c(11, 12, 13, 13, 14, 14, 14, 15, 16, 18, 18, 19)
```

Step 2: Executing the Two-Sample t-test using the t.test() Function

Once the data vectors (``group1`` and ``group2``) have been successfully loaded into the R environment, the analysis proceeds by invoking the built-in **t.test()** command. For a standard two-sample test comparing independent groups, the syntax is straightforward, requiring only the two vectors as arguments. By default, R performs a two-sided test and typically uses the Welch's modification (assuming unequal variances), unless otherwise specified.

The execution of this function generates a comprehensive output that summarizes the statistical findings. This output contains all the essential elements required for a complete interpretation: the calculated t-statistic, the associated degrees of freedom (df), the critical p-value, and the confidence interval for the true difference in means. Understanding how these components are derived and how they interact is key to sound statistical reporting.

The specific output below shows the results of comparing the height measurements of the two

plant species. The calculated values represent the empirical evidence we use to evaluate the initial premise that the mean heights are identical.

```
#perform two sample t-test
```

```
t.test(group1, group2)
```

Welch Two Sample t-test

data: group1 and group2

t = -2.5505, df = 20.488, p-value = 0.01884

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

95 percent confidence interval:

-5.6012568 -0.5654098

sample estimates:

mean of x mean of y

11.66667 14.75000

Deconstructing the R Output: Key Statistical Metrics

The output generated by R is highly structured, providing clear labels for each critical component of the statistical test. A detailed examination of each line is necessary to fully grasp the implication of the findings.

data: This field explicitly confirms the input used for the analysis. In this case, it indicates that the comparison was performed between the vectors designated as `group1` and `group2`. Verification of the data sources ensures that the correct variables were included in the comparison.

t: This value, **-2.5505**, is the computed t-statistic. The sign (negative) indicates that the mean of the second group (`group2`) is greater than the mean of the first group (`group1`). The absolute value reflects the standardized difference between the two sample means.

df: This represents the degrees of freedom, calculated here as **20.488**. Since R defaulted to the Welch Two Sample t-test (which does not assume equal variances), the degrees of freedom calculation uses the Satterthwaite approximation, resulting in a non-integer value. This value is crucial because it defines the specific t-distribution used to determine the critical regions and the associated p-value.

p-value: The p-value, found to be **0.01884**, is perhaps the most critical output element. It quantifies the likelihood of obtaining a t-statistic of -2.5505 or more extreme, given that the null hypothesis is true. This value directly informs the statistical decision.

alternative hypothesis: This line specifies the nature of the test conducted. In this case, the test was two-sided: the alternative hypothesis is that the true difference in population means is not equal to zero ($\mu_1 \neq \mu_2$). This is the standard setting for the function unless explicitly modified.

95 percent confidence interval: This range, , provides an estimated range of values for the true difference between the population means, based on the sample data. Since zero is not contained within this confidence interval, it strongly suggests that the true population means are indeed different, aligning with the conclusion drawn from the p-value.

sample estimates: These are the descriptive statistics for the input data. Specifically, the sample mean of group 1 was **11.667** and the sample mean of group 2 was **14.75**. These estimates quantify the observed difference in heights used to compute the t-statistic.

Making the Statistical Decision: Rejecting or Failing to Reject the Null Hypothesis

For this particular two-sample t-test, we established the formal statistical hypotheses as follows:

\$H_0\$: $\mu_1 = \mu_2$ (The mean heights of the two plant populations are equal.)

\$H_A\$: $\mu_1 \neq \mu_2$ (The mean heights of the two plant populations are *not* equal.)

The crucial step in the interpretation involves comparing the calculated p-value against the predetermined significance level, α . Using the standard $\alpha = 0.05$, we observe that the p-value of our test, which is **0.01884**, is substantially less than 0.05. This finding is statistically significant.

Since $P(0.01884) < \alpha(0.05)$, we must reject the null hypothesis (H_0). This rejection signifies that the observed difference between the sample means (11.667 vs. 14.75) is too large to be reasonably attributed to chance or random sampling error. Consequently, we conclude that there is sufficient statistical evidence to assert that the true mean height of plants between the two populations is genuinely different.

Advanced Usage and Customizing the `t.test()` Parameters

The `t.test()` function in R is highly flexible and allows users to modify several parameters to suit different testing requirements, such as directional hypotheses or dealing with paired samples. Understanding these parameters is essential for performing variations of the t-test beyond the default independent, two-sided test.

The general syntax for the `t.test()` function includes the following arguments, each controlling a

specific aspect of the hypothesis test setup and calculation:

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

where the parameters serve the following functions:

x, y: These are the names of the two numeric vectors that contain the data for the respective samples being compared.

alternative: This argument defines the nature of the alternative hypothesis (H_A). Options include **"two.sided"** ($\mu_1 \neq \mu_2$), **"less"** ($\mu_1 < \mu_2$), or **"greater"** ($\mu_1 > \mu_2$). Selecting a one-sided test changes how the p-value is calculated.

mu: This specifies the hypothesized value of the true difference in means under the null hypothesis. The default is $\mu = 0$, assuming no difference.

paired: Setting this to **TRUE** converts the function into a paired t-test, suitable for before-and-after measurements or matched pairs. The default is **FALSE** (independent samples).

var.equal: Setting this to **TRUE** forces the function to use the pooled variance estimate, conducting a classic Student's t-test. If set to **FALSE** (the default), the Welch's t-test is used, accommodating unequal variances and calculating fractional degrees of freedom.

conf.level: This determines the desired confidence interval for the true difference in means. The conventional standard is 0.95 (95%).

Summary of Interpretation Best Practices

In the example provided, we utilized the default settings of the **t.test()** function, leading to the following implied assumptions and parameters:

We conducted a two-sided alternative hypothesis test (H_A : $\mu_1 \neq \mu_2$).

We tested the null hypothesis that the true difference in means was exactly equal to zero.

We performed an independent two-sample t-test ($\text{paired} = \text{FALSE}$).

We did not make the strict assumption that the population variances were equal between the groups ($\text{var.equal} = \text{FALSE}$), resulting in the use of the Welch approximation.

We used a 95% confidence interval for estimating the difference.

To maximize the integrity and utility of your statistical findings, always confirm that your chosen parameters accurately reflect the underlying research design. If, for instance, you were comparing pre- and post-intervention scores on the same subjects, you would explicitly set $\text{paired} = \text{TRUE}$. Similarly, if strong theoretical or empirical evidence supports the assumption of equal variances, changing var.equal to **TRUE** would provide a slightly more powerful test, though the Welch test is generally safer.

Feel free to change any of these arguments when you conduct your own t-test, depending on the particular scientific question you aim to answer and the nature of the data collected. Rigorous interpretation depends not just on the output values but also on the correctness of the input parameters chosen.

ARABPSYCHOLOGY.COM