

How to Easily Perform a One Sample t-Test in SAS

Authored by
stats writer

December 1, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform a One Sample t-Test in SAS*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=103403>

Performing a one sample t-test within SAS is efficiently managed using the specialized statistical procedure, PROC TTEST. This powerful tool is designed to assess whether the mean of a single population differs significantly from a predefined, hypothesized value. To execute this procedure successfully, the user must meticulously define three critical components: the specific variable subject to testing, the theoretical or hypothesized mean (often denoted as μ_0), and the acceptable significance level (alpha, typically 0.05). Upon execution, PROC TTEST generates comprehensive output, detailing essential metrics such as the sample mean, the standard deviation, the degrees of freedom, the calculated t-statistic, and, most importantly, the corresponding p-value, which dictates the statistical conclusion.

The one sample t-test serves as a fundamental inferential statistical technique utilized specifically to determine whether the average value (the mean) of a population, estimated from a representative sample, is statistically equivalent to, or significantly different from, a certain target or hypothesized numerical value. This statistical comparison forms the basis of many research questions across various scientific disciplines.

This comprehensive tutorial is engineered to walk you through the precise methodology required to perform a robust and valid one sample t-test using the statistical analysis capabilities of SAS software.

Introduction to the One Sample t-Test and SAS

Before diving into the coding, it is essential to understand the underlying purpose of this test. The one sample t-test allows researchers to take a small, manageable sample from a larger, potentially infinite population and make an informed inference about that population's true average. For instance, if a manufacturer claims their product lasts 1000 hours, we can test a sample of products to see if their actual average lifespan is statistically different from that claim. This test is crucial when the population standard deviation is unknown, which is frequently the case in real-world data analysis.

The power of SAS lies in its procedural approach. Statistical operations are executed via dedicated procedures, known as PROCs. For mean comparisons, PROC TTEST is the standard command. This procedure not only calculates the t-statistic but also handles complex calculations related to degrees of freedom and confidence intervals, ensuring the researcher receives a full suite of necessary metrics for decision-making. We will be utilizing specific options within PROC TTEST to define our test parameters clearly, focusing on the hypothesis (H_0) and the desired significance level (α).

Prerequisites and Statistical Assumptions

For the results of a one sample t-test to be statistically valid and reliable, several key assumptions must be met. Ignoring these assumptions can lead to incorrect conclusions regarding the population mean. The primary assumption is that the sample observations must be independent of one another. In our example, the height of one plant should not influence the height measurement of any other plant in the sample.

Secondly, the population from which the sample is drawn should be approximately normally distributed. While the t-test is generally robust to minor violations of normality, especially with larger sample sizes ($n > 30$), it is a crucial consideration for smaller datasets like the one we will use. Analysts often perform initial graphical checks, such as histograms or Q-Q plots, before running the t-test, although PROC TTEST can also provide some normality diagnostics. Finally, the data should be measured on an interval or ratio scale, which is satisfied by continuous measurements such as height or weight.

Detailed Example Scenario: Botanical Research

Consider a practical research scenario often encountered in botany or agriculture. Suppose a botanist is investigating a specific species of exotic plant. Historical data or theoretical models suggest that the ideal or typical mean height for this mature species is exactly 15 inches. The botanist wishes to scientifically determine if the current population of plants she is studying exhibits a mean height that significantly deviates from this established 15-inch target.

To conduct this investigation, she meticulously collects a random sample of 12 mature plants from the population of interest. She records the precise height of each sampled plant in inches. The measured heights, which will form our dataset in SAS, are as follows: 14, 14, 16, 13, 12, 17, 15, 14, 15, 13, 15, and 14 inches.

Our objective is to employ the step-by-step methodology of the PROC TTEST in SAS to rigorously test the hypothesis. Specifically, we aim to discover if there is sufficient statistical evidence to conclude that the true population mean height for this species of plant is statistically different from the hypothesized value of 15 inches. This setup requires us to define both the null hypothesis (H_0) and the alternative hypotheses (H_A) based on the botanist's objective.

Step 1: Data Creation and Setup in SAS

The foundational step in any statistical analysis using SAS is structuring the raw data into a usable dataset. We begin by defining a new dataset, conventionally named `my_data`, and then using the `INPUT` statement to specify the variable name--in this case, `Height`. The `DATALINES` statement signals to SAS that the subsequent lines contain the raw data entries corresponding to

the input variable.

The following clean and well-structured code block illustrates the process of generating this plant height dataset within the SAS programming environment. It is crucial to ensure that the data is entered accurately, as any transcription error can skew the final statistical results. Following the data entry, we execute a simple `PROC PRINT` to verify the successful creation and accuracy of the dataset before proceeding to the actual hypothesis test.

```
/*create dataset*/  
data my_data;  
input Height;  
datalines;  
14  
14  
16  
13  
12  
17  
15  
14  
15  
13  
15  
14  
;  
run;  
  
/*print dataset*/  
proc print data=my_data;
```

Obs	Height
1	14
2	14
3	16
4	13
5	12
6	17
7	15
8	14
9	15
10	13
11	15
12	14

Upon successful execution of the data step and the print procedure, the table above confirms that our 12 observations have been correctly loaded into the `my_data` dataset. This confirmation ensures we have a solid foundation for the subsequent statistical analysis, mitigating any risk of performing calculations on incomplete or incorrectly structured input data.

Step 2: Executing the PROC TTEST Procedure

With the data prepared, we now move to the core of the analysis: utilizing `PROC TTEST` to perform the one sample t-test. This procedure requires several essential parameters to be specified directly in the command line to accurately define the nature of the test. The primary parameters include specifying the input dataset (`data=my_data`), defining the type of test (`sides=2` for a two-sided test, meaning we are checking if the mean is either greater than or less than the hypothesized value), setting the level of significance (`alpha=0.05`, which is the standard threshold), and most importantly, defining the null hypothesis value (`h0=15`, signifying that the population mean is hypothesized to be 15 inches).

Furthermore, the `VAR` statement within the procedure block is mandatory; it explicitly tells SAS which variable contains the measurements to be tested against the hypothesized mean. In this case, we specify `VAR Height`. The combined command structure below is the standard syntax for executing a one sample t-test that checks for a difference from 15 inches at the 5% significance level.

```
/*perform one sample t-test*/  
proc ttest data=my_data sides=2 alpha=0.05 h0=15;  
var Height;
```

run;

The TTEST Procedure					
Variable: Height					
N	Mean	Std Dev	Std Err	Minimum	Maximum
12	14.3333	1.3707	0.3957	12.0000	17.0000

Mean	95% CL Mean	Std Dev	95% CL Std Dev
14.3333	13.4624 15.2042	1.3707	0.9710 2.3273

DF	t Value	Pr > t
11	-1.68	0.1201

Once the code is successfully run, SAS generates a series of output tables. These tables contain all the descriptive and inferential statistics necessary for drawing a formal conclusion regarding the null hypothesis. The interpretation of these tables is the next crucial phase of the analysis.

Interpreting the Descriptive Statistics Output

The first table produced by PROC TTEST is the descriptive statistics summary. This table provides a snapshot of the characteristics of the sample data collected, helping the analyst confirm the sample size and gauge the variability and central tendency before looking at the inferential results. Understanding these statistics is vital for contextualizing the subsequent t-test results.

Key metrics displayed in this initial output include:

N (Total Observations): This is the sample size, confirming that **12** plants were included in the test.

Mean (Sample Mean): The arithmetic average of the observed heights is calculated as **14.3333** inches. This observed mean is slightly lower than the hypothesized value of 15.

Std Dev (Sample Standard Deviation): This measures the spread or variability of the data points around the mean, calculated here as **1.3707**. A smaller value indicates less variation in plant heights.

Std Error (Standard Error of the Mean): Calculated as the sample standard deviation divided by the square root of the sample size ($\text{\$/sqrt}\{n\}$), the standard error is **0.3957**. This value represents the estimated standard deviation of the sampling distribution of the mean.

Minimum and Maximum: These values provide the range of the observed data, with a minimum

height of **12** inches and a maximum height of **17** inches.

Following the descriptive statistics, the output provides a **95% Confidence Interval (C.I.)** for the true population mean (μ). This interval is crucial because it gives us a range of plausible values for the true average height of all plants of this species, based on our sample data.

95% C.I. for μ : The interval runs from .

Since the hypothesized value of 15 inches falls within this 95% confidence interval, this provides initial evidence, even before reviewing the p-value, that the null hypothesis may not be rejected. If the hypothesized mean fell outside this interval, we would typically reject H_0 .

Analyzing the p-value and Drawing Conclusions

The final and most critical table in the PROC TTEST output presents the actual t-test statistics, which quantify the difference between the sample mean and the hypothesized population mean in standard deviation units. This table contains the t-statistic and the associated p-value, which is the probability of observing our sample data (or more extreme data) if the null hypothesis were true.

t Test Statistic: Calculated as **-1.68**. This negative value indicates that the sample mean (14.3333) is below the hypothesized mean (15).

p-value: The two-sided p-value for this test is **0.1201**.

Note: The t test statistic calculation is defined by the formula:

$$t \text{ test statistic} = (\bar{x} - \mu) / (s / \sqrt{n})$$

$$t \text{ test statistic} = (14.3333 - 15) / (1.3707 / \sqrt{12})$$

t test statistic = -1.68, confirming the automated SAS calculation.

To make a formal decision, we must explicitly state the hypotheses being tested:

H0 (Null Hypothesis): $\mu = 15$ inches (The true mean height is 15 inches.)

HA (Alternative Hypothesis): $\mu \neq 15$ inches (The true mean height is different from 15 inches.)

The crucial step in interpretation is comparing the calculated p-value (0.1201) to the pre-specified significance level, or alpha ($\alpha = 0.05$). The decision rule dictates that if the p-value is less than α , we reject the null hypothesis. Conversely, if the p-value is greater than α , we fail to reject the null hypothesis. Since the p-value of **0.1201** is clearly greater than 0.05, we must **fail to reject the null hypothesis**.

This statistical outcome leads to the conclusion that we do not possess sufficient statistical

evidence, based on this sample of 12 plants, to confidently assert that the true average height of this species is significantly different from 15 inches. The observed difference between the sample mean (14.3333) and the hypothesized mean (15) is likely attributable to random sampling variability, rather than a genuine difference in the population mean height.

Further Statistical Analyses in SAS

Mastering the PROC TTEST for a one sample t-test provides a strong foundation for conducting more complex inferential statistics in SAS. Depending on the research question, analysts often proceed to other specialized procedures, such as two-sample t-tests (comparing two population means) or paired t-tests (comparing means before and after an intervention).

The following tutorials explain how to perform other common statistical tests in SAS, building upon the skills developed here: