Learning the Paired Samples t-test: A Step-by-Step Guide Using the TI-84 Calculator

Authored by

Mohammed looti

November 8, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Learning the Paired Samples t-test: A Step-by-Step Guide Using the TI-84 Calculator*. PSYCHOLOGICAL STATISTICS. Retrieved from https://statistics.arabpsychology.com/?p=13303

The Paired Samples t-test is a fundamental procedure in inferential statistics, specifically designed to compare the **means** of two groups that are inherently related or dependent. This test is essential when data consists of dependent samples, meaning that every single observation in the first sample is directly matched or paired with a corresponding observation in the second sample. Classic examples of its application include rigorous before-and-after experiments, crossover studies, or research involving carefully matched subjects to control for individual variability.

This comprehensive guide provides detailed, expert instructions on the most accurate and efficient method for conducting a paired t-test using the powerful functionalities available on the **TI-84 calculator**. By transforming the dependent measurements into a single variable of difference scores, the TI-84 allows students and researchers to rapidly determine if a statistically significant effect or change exists between the two paired conditions. Mastering this technique is crucial for accurate data analysis in many scientific fields.

The Essential Logic of the Paired Samples t-test

The paired t-test--also frequently termed a dependent samples t-test or a paired difference test--is conceptually distinct from the independent samples t-test. The crucial element of dependence stems from the fact that the data points are not drawn randomly from two entirely separate populations; rather, they represent two measurements (e.g., pre-treatment and post-treatment) taken from the exact same individual or unit under different, controlled conditions. This sophisticated experimental design is highly valuable because it inherently minimizes the random noise or variability often caused by individual differences, thereby making the test significantly more powerful for successfully detecting genuine effects.

The fundamental mathematical principle of the paired t-test is to simplify the comparison by converting the two dependent variables into a single, highly informative variable: the difference score. If there is genuinely no effect or difference existing between the two conditions being compared, then the true population mean of these difference scores should theoretically be zero ($mu_D = 0$). The statistical test then rigorously assesses whether the observed mean difference calculated from the sample is sufficiently far removed from this hypothesized value of zero to be categorized as statistically significant.

Executing this specialized test on a computational tool like the **TI-84 calculator** requires creatively utilizing its standard T-Test function, but applying it exclusively to the newly generated column of difference scores. This methodical approach dramatically streamlines the traditional manual calculation process. It allows both researchers and students to dedicate their focus to the critical interpretation of the resulting test statistics and calculated p-values, which are the cornerstone for drawing informed, data-driven conclusions regarding the underlying population parameters.

Defining the Research Problem and Collecting Paired Data

To clearly illustrate the application of this procedure, let us consider a practical research scenario. A team of automotive engineers is interested in rigorously evaluating the true impact of a newly developed fuel treatment additive on the average miles per gallon (MPG) efficiency of a popular sedan model. To ensure the highest level of experimental control and rigor, they devise a paired experiment where a sample of 11 distinct cars are tested exactly twice: first, under standard driving conditions (the baseline measurement, without the treatment), and second, immediately after the new fuel treatment has been systematically applied.

Since every single car effectively serves as its own internal control--the measurement taken before the treatment is intrinsically linked to the measurement taken after the treatment--the collected data are fundamentally paired. This specific experimental design absolutely mandates the use of a **paired t-test**, as this method is uniquely suited to accurately account for the dependency that exists between the 'MPG without treatment' and the 'MPG with treatment' measurements. If the new fuel treatment is genuinely effective, the analysis should reveal a statistically consistent, non-zero difference in MPG across all 11 tested vehicles.

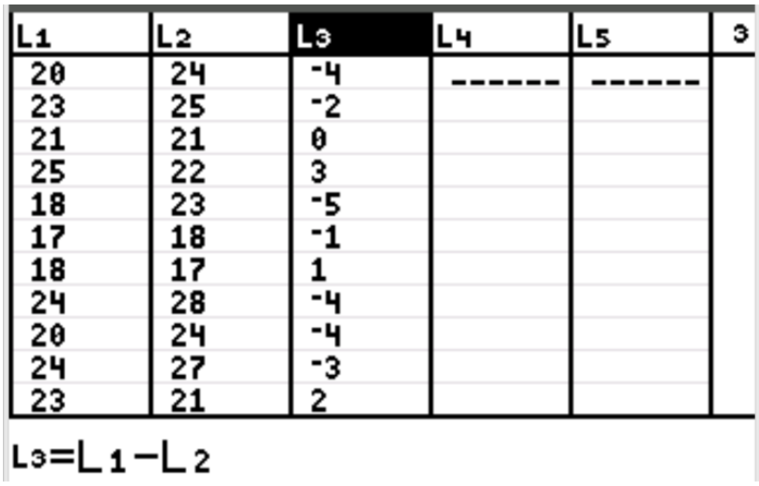
The formal structure of the hypothesis test is established as follows: The null hypothesis ($text{H}_0$) posits that the true mean difference in MPG is exactly zero ($mu_D = 0$), implying the treatment has exerted no measurable effect. Conversely, the alternative hypothesis ($text{H}_text{a}$) asserts that the true mean difference is not zero ($mu_D neq 0$), suggesting that the treatment has caused a significant change in MPG. The subsequent steps will meticulously detail how to use the TI-84 to perform the calculation necessary to test these critical assumptions.

Step 1: Data Entry and Calculating the Difference Scores

The required first step for successfully performing this analysis is the precise and accurate input of the raw paired data into the calculator's dedicated statistical lists. Begin this process by pressing the Stat button, and then immediately selecting the EDIT option from the menu that appears. In List 1 (L1), meticulously enter all the observed MPG values for the **control group** (the measurements taken without the fuel treatment). Subsequently, enter the corresponding paired MPG values for the **treatment group** (measurements taken with the fuel treatment) into List 2 (L2). It is paramount that the paired observations align horizontally across L1 and L2.

The most critical procedural step for the paired t-test is the calculation of the difference between the two paired measurements for each car, which must be stored in List 3 (L3). To efficiently generate these essential difference scores, scroll the cursor up to the very top of the list columns until the header **L3** is clearly highlighted. Once L3 is highlighted, you can input the formula that directs the calculator to perform the subtraction. Press 2nd and then 1 to recall L1, followed by the standard minus sign ($-$), and then press 2nd and 2 to recall L2. The complete formula displayed should read exactly as: **L1 - L2**. Conclude this step by pressing Enter.

Upon execution, the TI-84 will instantaneously populate every cell in L3 with the corresponding calculated difference score between the L1 and L2 values. This newly generated List 3 now contains the required data points--the paired differences--which serve as the sole dataset for the subsequent statistical t-test calculation. Always visually verify that your data entry and the automatically calculated differences match the visual representation provided below before proceeding to the testing phase.



L1	L2	L3	L4	L5	3
20	24	-4	------	------	
23	25	-2			
21	21	0			
25	22	3			
18	23	-5			
17	18	-1			
18	17	1			
24	28	-4			
20	24	-4			
24	27	-3			
23	21	2			

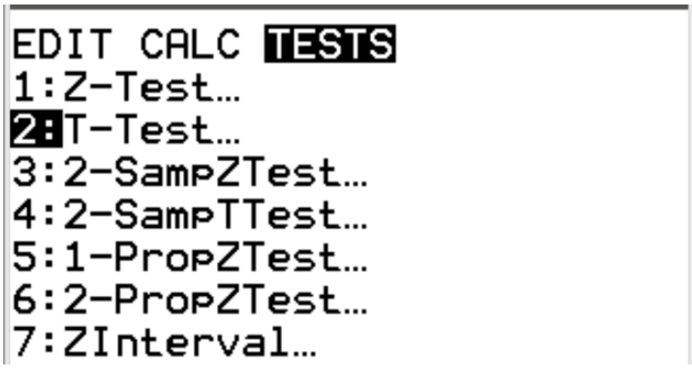
L3=L1−L2

Step 2: Configuring and Running the One-Sample T-Test on L3

To formally execute the paired t-test, we must now configure the calculator to run a standard T-Test, directing it exclusively at the difference scores that are stored in List 3. This crucial procedural substitution effectively converts the paired test into a one-sample t-test on the differences. Begin by pressing Stat once more, then scroll across the menu ribbon to the **TESTS** category, and finally scroll down to select option **2:T-Test**. Press ENTER to open the configuration screen for the T-Test.

The calculator requires specific parameter inputs to perform the analysis correctly. Since we are using the raw difference scores already computed and stored in L3, we must select **Data** as the primary input method (as opposed to Stats). We must also carefully specify the value for mu_0, which represents the hypothesized mean difference under the null hypothesis. For virtually all paired t-tests, the hypothesis is that there is no difference, thus we must accurately enter **0** for mu_0 and press ENTER.

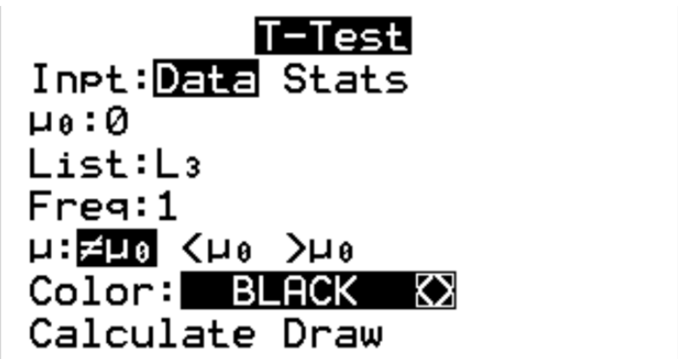
Next, specify the list containing the difference values by entering **L3** for the **List** parameter (achieved by pressing 2nd and then 3). The **Freq** parameter should always be left as the default value of **1** for raw data input. The final, yet vital, choice is selecting the form of the alternative hypothesis (mu). Since our initial research question asked if the fuel treatment caused a general **change** (implying a two-sided test), we must highlight $mathbf{neqmu_0}$ and confirm this selection by pressing ENTER. This choice mathematically confirms our hypothesis that the true mean difference is simply not equal to zero. Note that if the research had been directional (e.g., hypothesizing a specific MPG increase), we would choose the appropriate one-tailed option ($<mu_0$ or $>mu_0$).



After all parameters have been correctly set, scroll the cursor down to highlight the option labeled **Calculate** and press ENTER. The **TI-84 calculator** will then process the dataset in L3 and immediately display the calculated statistical results.

Step 3: Analyzing and Interpreting the Statistical Output

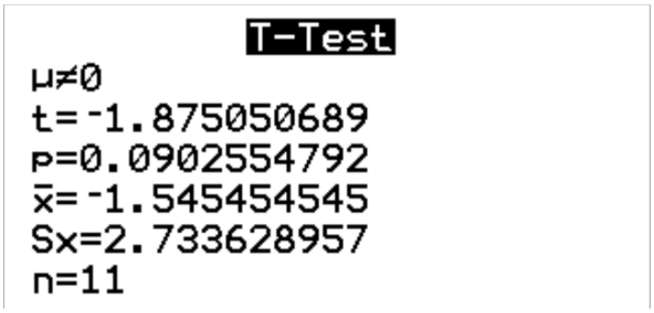
Upon completion of the calculation, the TI-84 screen will present a concise summary of the one-sample t-test results, derived entirely from the difference scores stored in L3. Developing a clear understanding of what each output variable represents is absolutely essential for accurately drawing conclusions about the fuel treatment's hypothesized effect on vehicle MPG.



The primary output screen provides the following critical statistics:

$mu neq 0$: This line serves as a confirmation of the two-sided alternative hypothesis that was

utilized in the test setup (that the true mean difference is not zero).
t = -1.8751: This value is the calculated t test-statistic. This metric quantifies the number of estimated standard errors that the sample mean difference lies away from the hypothesized population mean difference of zero. The negative sign, in this case, simply reflects that the average difference (L1 - L2) resulted in a negative value.
p = 0.0903: This is the critical p-value corresponding directly to the calculated t-statistic. It represents the conditional probability of observing a sample mean difference as extreme as -1.5455 (or even more extreme) if the null hypothesis were entirely true.
$\overline{x} = -1.5455$. This represents the sample mean difference, which is computed as the arithmetic average of all the L3 values (Group 1 - Group 2). Practically, this indicates that, on average, the MPG decreased by 1.5455 units after the treatment was applied.
$s_{x} = 2.7336$. This is the standard deviation of the differences (the standard deviation of L3). It provides a measure of the dispersion or variability of the individual difference scores around the calculated mean difference.
n = 11: This figure denotes the total number of paired samples, or observations, that were used in the formal hypothesis test.



Drawing Statistically Valid Conclusions

The conclusive phase of any formal hypothesis test involves comparing the calculated p-value against the researcher's predetermined threshold of statistical significance, which is conventionally known as the alpha level (α). In the vast majority of scientific and academic research, the alpha level is fixed at 0.05 (or 5%). The fundamental decision rule governing this comparison is clear and unambiguous: if the calculated p-value is strictly less than α, we must reject the null hypothesis; conversely, if the p-value is greater than α, we must fail to reject the null hypothesis.

Analyzing the results from this specific scenario, the calculated p-value is 0.0903. Since 0.0903 is observably **greater than** the standard significance level of 0.05, the statistical imperative is to **fail to reject the null hypothesis**. This outcome--the failure to reject--is the mathematical indication that the observed mean difference of -1.5455 MPG is not statistically large enough, when accounting for the sample variability and the relatively small sample size, to be deemed statistically significant at the 5% level of confidence.

In terms of practical interpretation, based on the rigorous standards of this paired t-test analysis, the researchers cannot claim to possess sufficient statistical evidence to definitively conclude that the new fuel treatment causes a significant change in the average MPG of the tested vehicles. Although the sample data showed a minor average decrease, the inherent variation observed among the 11 cars suggests that this difference could plausibly be attributed to natural, random sampling variability or chance, rather than representing a true, systematic effect caused by the fuel treatment itself.