

Perform a Paired Samples t-Test in SAS

Authored by
Mohammed looti

November 1, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Perform a Paired Samples t-Test in SAS*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=7625>

A [Paired Samples t-Test](#), frequently known as the dependent samples t-test, is a cornerstone statistical procedure utilized when researchers aim to compare the [means](#) derived from two groups where the observations are intrinsically linked or paired. This test is indispensable in experimental designs where measurements are taken from the same subjects under different conditions, such as classic pre-test/post-test evaluations, or in studies involving rigorously matched pairs like siblings or specific geographical locations.

The primary advantage of employing the Paired Samples t-Test is its enhanced statistical power, achieved by effectively controlling for the inherent variability that exists between individual subjects. By focusing the analysis on the difference between the paired measurements, rather than comparing the overall group averages independently, we isolate the effect of the intervention. This detailed tutorial provides a step-by-step methodology for executing and rigorously interpreting a paired samples t-test using the **Statistical Analysis System (SAS)**, a robust and powerful software suite essential for advanced statistical modeling and complex data management.

The Statistical Framework of the Paired Samples Design

It is vital to distinguish the Paired Samples t-Test from its counterpart, the Independent Samples t-Test. The structure of paired data inherently implies dependency; the scores or measurements within one sample are directly related to the scores in the second sample, as they originate from the identical subject or from a subject specifically matched to another. This dependency is the defining characteristic that mandates the use of this specialized test, which calculates the mean of the differences for each pair, rather than simply assessing the difference between the two overall group means.

The central objective of this analysis is to rigorously determine whether the average difference observed between the paired observations deviates significantly from zero. A statistically significant non-zero average difference provides compelling evidence that the treatment, intervention, or time lapse occurring between the two measurements resulted in a measurable and reliable effect. If the difference is not significantly different from zero, we conclude that the intervention likely had no effect.

Before proceeding with the formal test execution, the validity of the results rests upon fulfilling specific statistical assumptions. The most critical assumption is that the differences between the paired observations must be approximately **normally distributed**. Furthermore, it is assumed that the observations within each pair are dependent, while the differences between pairs are independent. If these distributional assumptions are severely violated--particularly in cases involving small sample sizes--researchers must consider employing non-parametric alternatives, such as the Wilcoxon Signed-Rank Test, to ensure the reliability of their conclusions.

Case Study: Evaluating Educational Program Effectiveness

To illustrate the practical application of the Paired Samples t-Test, consider a typical educational scenario. A university professor has developed an intensive study program aimed at boosting student scores in a particularly challenging core academic subject. The professor must design a rigorous evaluation method to quantify the program's efficacy. A paired design offers the most appropriate methodology for this task.

The methodology involves recruiting 15 student volunteers. Initially, these students participate in a standardized baseline assessment, known as the **pre-test**, to establish their initial level of competency. Following this, all participants are required to engage fully with the intensive study program for a predefined duration of one month. Finally, they complete a subsequent assessment--the **post-test**--which is carefully constructed to possess similar difficulty and scope as the initial pre-test.

Because every student contributes two data points--both a pre-test score and a post-test score--their individual data is intrinsically linked. This dependency structure confirms the necessity of using the [Paired Samples t-Test](#) to accurately assess the change in performance attributable to the study program. The scores collected for each of the 15 participants, structured for analysis, are displayed visually in the image below:

Student	Pre-Test Score	Post-Test Score
1	88	91
2	82	84
3	84	88
4	93	90
5	75	79
6	78	80
7	84	88
8	87	90
9	95	90
10	91	96
11	83	88
12	89	89
13	77	81
14	68	74
15	91	92

The core statistical inquiry we must address is whether the average post-test score is significantly

higher or lower than the average pre-test score, thereby providing statistical evidence that the intensive study program had a tangible impact. We will now proceed through the necessary steps within the powerful [SAS](#) environment to answer this research question definitively.

SAS Implementation: Step 1 - Data Preparation and Structure

The foundational step in initiating any statistical analysis in SAS involves defining the dataset and ensuring its accurate input and structure. For the Paired Samples t-Test, the data must be structured in a wide format, meaning the paired measurements (pre-test and post-test scores) are organized into separate, corresponding variables within the same observation row.

We begin by utilizing the **DATA step** to formally create a dataset, which we name `test_scores`. The subsequent **INPUT statement** explicitly defines the two numerical variables: `pre` (pre-test score) and `post` (post-test score). The raw data is then entered efficiently using the `datalines` statement, meticulously ensuring that each line of input represents the paired scores belonging to a single student participant.

The following [SAS](#) code block not only accomplishes the crucial task of data creation but also includes a subsequent procedure to display an initial view of the dataset for verification purposes:

```
/*create dataset for 15 students*/
```

```
data test_scores;
```

```
input pre post;
```

```
datalines;
```

```
88 91
```

```
82 84
```

```
84 88
```

```
93 90
```

```
75 79
```

```
78 80
```

```
84 88
```

```
87 90
```

```
95 90
```

```
91 96
```

```
83 88
```

```
89 89
```

```
77 81
```

```
68 74
```

```
91 92
```

```
;
```

```
run;
```

```
/*view dataset to confirm data entry accuracy*/  
proc print data=test_scores;
```

Executing the proc print command is a necessary step to confirm the successful and accurate loading of the data into the SAS environment. The resulting output confirms that the scores match the original data table, ensuring the integrity of the data prior to statistical processing:

Obs	pre	post
1	88	91
2	82	84
3	84	88
4	93	90
5	75	79
6	78	80
7	84	88
8	87	90
9	95	90
10	91	96
11	83	88
12	89	89
13	77	81
14	68	74
15	91	92

With this preliminary check complete, the dataset is fully prepared and structured correctly for the subsequent statistical analysis step.

SAS Implementation: Step 2 - Executing the Paired Samples t-test

The formal execution of the Paired Samples t-Test in [SAS](#) is carried out using the highly versatile **PROC TTEST** procedure. This procedure is capable of handling various t-test types, but for paired data, a specific syntax is required to inform SAS of the dependent relationship between the variables.

Crucially, we employ the **PAIRED** statement. The syntax `paired pre*post` explicitly instructs SAS to calculate the difference score for every single observation by subtracting the pre score from the post score (or vice versa, depending on the order specified). The procedure then proceeds to test

whether the [mean](#) of these resulting difference scores is statistically significantly non-zero. Additionally, the option `alpha=.05` is specified to set the desired significance level, which is standard practice in hypothesis testing.

The necessary code required to execute this robust statistical analysis is remarkably concise, demonstrating the efficiency of the SAS language:

```
/*perform paired samples t-test using PROC TTEST*/  
proc ttest data=test_scores alpha=.05;  
paired pre*post;  
run;
```

Upon successful execution of the PROC TTEST command, SAS generates a comprehensive output report. This report is structured to include detailed descriptive statistics for both variables, the calculated confidence intervals for the mean difference, and the essential inferential test results required to make a definitive statistical decision regarding the hypothesis.

Interpreting the Output and Drawing Conclusions

The output produced by `proc ttest` contains all the necessary metrics to evaluate our research hypotheses within the framework of statistical significance testing. Our decision-making process is fundamentally guided by the definitions of the [null hypothesis](#) (H_0) and the alternative hypothesis (H_1).

For this specific study evaluating the intensive program, the hypotheses are formally stated as follows:

H_0 (Null Hypothesis): The true mean difference between pre-test and post-test scores is exactly zero. Statistically, this implies the study program had no measurable effect on student performance.

H_1 (Alternative Hypothesis): The true mean difference between pre-test and post-test scores is *not* zero. This suggests that the study program successfully induced a change (either positive or negative) in student scores.

The TTEST Procedure

Difference: pre - post

N	Mean	Std Dev	Std Err	Minimum	Maximum
15	-2.3333	3.0394	0.7848	-6.0000	5.0000

Mean	95% CL Mean	Std Dev	95% CL Std Dev
-2.3333	-4.0165 -0.6502	3.0394	2.2252 4.7935

DF	t Value	Pr > t
14	-2.97	0.0101

When analyzing the core SAS output table (the "T-Tests" section), two key areas must be examined: the descriptive statistics of the difference and the inferential test results:

Mean Difference: The calculated average difference (Post score minus Pre score) is reported as **-2.3333**. Since we subtracted the Pre score from the Post score, this negative value indicates that, on average, students scored 2.33 points higher on the post-test after completing the program.

95% [Confidence Interval \(CI\)](#) for Mean difference: This critical interval is calculated as . The CI provides a range of plausible values for the true population mean difference. Because this interval does not encompass the value zero, we gain initial, strong evidence suggesting that the true difference is not zero.

The formal decision regarding the [null hypothesis](#) relies on the t-statistic and its associated [p-value](#):

t Test Statistic: The standardized test value is **-2.97**. This statistic quantifies the magnitude of the observed difference relative to the variability (standard error) within the sample.

Two-sided [p-value](#): This value is reported as **.0101**. This represents the probability of observing a mean difference as extreme as -2.3333 (or more) if the null hypothesis were true and no effect existed.

Since the calculated p-value (**0.0101**) is substantially smaller than our pre-determined significance level ($\alpha = 0.05$), we meet the established criterion for achieving statistical significance. Consequently, we must firmly reject the null hypothesis. This rejection allows us to conclude that there is sufficient statistical evidence to assert that the true mean test score differs significantly for students before and after their participation in the intensive study program. Given the negative mean difference (indicating Post > Pre), we confidently conclude that the program had a positive and statistically significant impact on student performance.

Summary and Further Considerations

The Paired Samples t-Test stands as an exceptionally valuable statistical tool for researchers evaluating interventions where measuring change within the same subject is paramount. Our analysis, efficiently executed using [SAS](#), clearly demonstrated a statistically significant improvement in the test [means](#) following the implementation of the study program, confirming its efficacy.

While the t-test provided a definitive statistical conclusion, ethical and methodological rigor demands that researchers always verify the underlying assumptions of the test. Specifically, the assumption regarding the normality of the difference scores must be checked. SAS facilitates this crucial step by providing optional output, such as Q-Q plots and formal normality tests (e.g., Shapiro-Wilk test), which should be reviewed to confirm the robustness of the t-test results. If the sample size is extremely small, or if the distribution of the differences is highly non-normal, a non-parametric alternative, such as the Wilcoxon Signed-Rank Test, would be the more statistically appropriate choice.

This methodology is highly versatile and applicable across numerous disciplines, including clinical trials (e.g., comparing a patient's biomarker levels before and after receiving a drug) and quality control engineering (e.g., comparing instrument accuracy between two different calibration methods).

Additional Resources for Statistical Analysis in SAS

To broaden your foundation in statistical computing, the following tutorials explain how to perform other common statistical tests using SAS, providing essential skills for diverse data analyses: