

Continue



Paired sample t-test

Medical example dataTest Score example dataThe dependent samples T-test (or paired samples t-test) is a statistical test that determines whether there is a difference between two dependent groups or samples. The dependent samples T-test, or also known as the dependent t-test, tests whether the mean values of two dependent groups differ significantly from each other. Why do you need the dependent t-Test? You need the paired t-test whenever you survey the same group or sample at two points in time. For example, you might be interested in whether a rehabilitation program has a positive effect on physical fitness. Since you can't ask all the people who go to rehab, you use a random sample. You can then use the paired t-test to infer the population from the sample. What are dependent or paired samples? In dependent samples, these measured values are available in pairs. The pairs result from repeated measurements, parallelization or matching. This can be the case, for example, in longitudinal studies with several measurement points (time series analyses) or in intervention studies with experimental designs (before-after measurement). An example of dependent sampling is when the weight of a group of people is measured at two points in time. A person can then be uniquely assigned a weight at the first and second measurement time points and the difference in the measured values can be calculated in each case. If more than two measurement times are available, ANOVA with repeated measures is used. What is the advantage of a dependent t-test over an independent t-test? The question of whether to use a dependent t-test or an independent t-test is, of course, already determined as part of the study design, and it is not possible to arbitrarily use either one test or the other. Therefore, the question is rather which type of study makes more sense: Conducting a study with one group of participants who are measured twice. To conduct a study with two separate groups of participants, each measured once. The major advantage of a repeated-measures design is that then uses the paired t-test is that individual differences between participants can be eliminated. This means that the probability of detecting a (statistically significant) difference, if one exists, is higher with the paired t-test than with the independent t-test. Example of a paired sample t-test The t-test for dependent samples has numerous applications, here are three examples. You want to check whether a new drug increases memory performance. You test the memory performance of 40 people before and after they take the medicine. A screw factory complains about very high downtimes at its 5 production plants. You are now to find out whether a newly introduced lubricant has an influence on the downtimes. For this you compare the downtimes of the 5 plants before and after the introduction of the new lubricant. You want to find out if there is a change between 2010 and 2015 in terms of health consciousness of the German population. For example, you could do this rely on data from the Socio-Economic Panel (SOEP). The SOEP is a representative repeat survey of private households in Germany. The survey will include always asks the same people at regular intervals about the same topics. To get your question to answer you compare the health consciousness of the persons in the year 2010 and 2015. Medical example dataTest Score example data Research Question and hypotheses In order to be able to calculate a t-test for dependent samples, you first need to define a question and the hypotheses In a t-test for dependent samples, the general question is: Is there a statistically significant difference between the mean value of two dependent groups? The questions for the above examples arise as follows: Does the new drug help to increase memory performance? Does the newly introduced lubricant have an influence on downtimes? Has the health consciousness of the German population changed between 2010 and 2015? Now the hypothesis can be derived from the question. In the hypothesis, a preliminary, i.e. unsubstantiated, assumption is made which is to be tested. In the case of a t test for dependent samples, the hypotheses are: Null hypothesis H0: The mean value of the two dependent groups is equal. Alternative hypothesis H1: The mean values of the two dependent groups are different. Assumptions paired t-Test Of course, the assumptions must be checked before calculating the paired t-test. If the prerequisites 2 and 3 (listed below) are not fulfilled, the Wilcoxon test must be used. The Wilcoxon test is the non-parametric counterpart of the paired t-test. As the name paired t-test already suggests the groups must be dependent, i.e. a value of one group must relate to a value of the other group. The weight of one and the same person is measured before and after a diet. Researchers measure the weight of people who have been on a diet and people who have not. 2. The variables are interval scaled In the t-test for dependent samples, the difference between the two dependent values is calculated and then the mean value. This only makes sense if the values are metric The salary of a person (in Euro) The educational level of a person 3. The differences of the paired values are normally distributed. The difference between the paired values must be normally distributed. The difference from the weight of one person at two points in time. The difference in the number of points after throwing two dice. How does a paired t-test work? In the paired t-test, the difference is calculated from each paired case. The mean value is then calculated from these differences. Depending on how large the mean value is and how large the standard error of the mean value is, a statement is then made as to how likely it is that this result has arisen by chance. Calculate t-test for paired samples For the calculation of the t-test for dependent samples, the difference of each pair from the two groups is first formed. From the resulting differences, the mean value $\bar{x}$ diff is then calculated. The calculation of the test statistics t is now equal to the t test for one sample. If there is no difference between the two groups, the mean value of the difference $\bar{x}$ diff is zero. So the question is, is there a difference between $\bar{x}$ diff and zero. The test statistic t for the t-test for dependent samples is then calculated as where is the standard error of the mean value = Difference between the groups = Mean value of the difference between the two groups = Sample size = Standard deviation = Estimated standard error of mean Effect size dependent t-test The indication of the effect size is very important for empirical studies. To make a statement about the effect size in a t-test for dependent samples, the following equation can be used In general, it can be said about the effect size: Effect size d: 0.2 small Effect Effect size d: 0.5 medium Effect Effect size d: 0.8 large Effect Calculate with DATAtab In the paired t-test example it is examined whether the summer holidays have an effect on the physical fitness of statistics students. So the question is: Do the summer holidays have an effect on the physical fitness of statistics students? For this purpose, a fitness test is carried out once before and once after the holidays for 10 statistics students (2 measurement points). The average difference of the measured value pairs (before and after the holidays) is zero. The semester break has no influence on the physical fitness of the students. Since two test results always come from one student, there is a dependency between the two samples. Therefore the paired t-test is used. Statistics student Points before holidays Points after holidays 1 60 61 2 70 71 3 40 38 4 41 39 5 40 38 6 40 33 7 45 55 8 48 56 9 30 38 10 50 68 After copying the upper table into the dependent t-test calculator you can calculate the t-test. The results are as follows: Statistics n Mean value Standard deviation Standard error of the mean Points before holidays 10 46.4 11.452 3.622 Points after holidays 10 49.7 14.095 4.457 Correlation n Correlation before holidays - after holidays 10 0,847 Paired t-test t df p before holidays - after holidays -1.39 9 0,197 95% confidence interval of the difference Mean Standard deviation Standard error of the mean Lower Upper Points after holidays -3.3 7.5 2.37 -8.66 2.06 Interpret t-test for dependent samples If the calculated p-value is smaller than the specified significance level (usually 5%), the null hypothesis is rejected, otherwise it is not rejected. For the upper example, you can report the results as follows: The score of the variable before the vacations had lower values (M = 46.4, SD = 11.452) than the score of the variable after the vacations (M = 49.7, SD = 14.095). A t-test for dependent samples showed that this difference was not statistically significant, t(9) = -1.392 p = .197, 95% confidence interval [-8.664, 2.064]. This results in a p-value of 0.197 which is above the defined significance level of 0.05. The t-test result is therefore not significant and the null hypothesis is not rejected, many illustrative examples ideal for exams and theses statistics made easy on 454 pages 6th revised edition (March 2025) Only 8.99 € Free sample "It could not be simpler" "So many helpful examples" Paired sample t-tests are a commonly used statistical procedure used to compare two populations that are related in some way. They are often used for comparing dependent groups, such as the before and after results of an experiment. Data scientists must have a thorough understanding of the concept of paired sample t-test in order to craft accurate and reliable results when analyzing data. In this blog post, we will explore the formula, assumptions, and examples of paired sample t-tests. What is Paired Sample T-Test and Why is it needed? Paired sample t-tests are used to test whether means of same or similar group different from each other under separate conditions (before and after intervention). For example, you would want to test the efficacy of a drug on the same group of patients before and after drug is given to the patients. When the goal is to assess the effect of a particular treatment before and after, you can use paired sample t-tests. Here are some real-world examples where paired sample t-test can be used: Testing the effectiveness of a new advertising campaign compared to the existing one. Marketing teams will be interested in this analysis. Examining if there is an improvement in website performance with the introduction of a new feature. This will be very useful for product managers to evaluate impact of the new product features & functionality. Measuring the effectiveness of employee training programs by comparing pre and post test results. Evaluating the impact of new packaging design on consumer purchasing decisions. Comparing the average price of a product before and after a price change to measure the effect of the change on sales. Here is going to be the null and alternate hypothesis for paired sample t-tests: Null Hypothesis: The mean of the difference of change variable between the two paired samples is zero (i.e. there is no statistically significant difference between them). In other words, the impact of treatment provided to group showed no significant difference from the earlier state. Alternate Hypothesis: The mean of the difference of variable between the two paired samples is not zero (i.e. there is a statistically significant difference between them). In other words, the impact of treatment provided to group showed significant difference from the earlier state. T-Statistics Formula for Paired Sample T-Test The following represents the formula for T-statistics: $T = \frac{\text{(mean difference between pairs)} / \text{(standard deviation of differences)} / \sqrt{\text{sample size}}}{\text{square root of sample size}}$ Where: Mean difference between pairs = (sum of differences between pairs) / sample size Standard deviation of differences = square root of [summation of variance of differences between pairs / (sample size - 1)] Here is how the above would look like: In the above formula, $\frac{d}{\sqrt{n}}$ is the mean of difference in the change variable, $\frac{d}{\sqrt{n}}$ is standard deviation of difference in change variable and n is the size of the sample. Paired Sample T-Test Example Based on the above formula, lets take a look at the example. Lets say the hypothesis is that walking for 1 hour a day for 3 months result in weight loss. In the example below, 10 people were subjected to the experiment and their weights (in KGs) were recorded after 3 months. Null hypothesis is that the mean of difference between 10 pairs is zero. In other words, the change is weight is not significant. Alternate hypothesis is that the mean of difference between 10 pairs is less than zero. In other words, the weights decrease after the walking is significant based on the evidence. Here is the calculation In the above test, the t-statistics is greater than critical t-value. In other words, it does not lie in the critical region. Thus, we don't have enough evidence to reject the null hypothesis. So, based on the evidence, it can not be claimed that walking for 1 hours every day for 3 months would result in weight reduction. Assumptions of Paired Sample T-Tests Before conducting a paired sample t-test there are certain assumptions you must make about your data set. The observations within the two samples must be related or paired in a meaningful way. This means that each observation from one sample should correspond to an observation in the other sample, typically with a causal relationship between them. For example, if you are measuring customer satisfaction before and after implementing a new policy, the 2 samples consist of ratings given by same customers before and after implementation. All observations must be independent of each other. This means that no two pairs of observations should have any association with each other – they should all be measured independently and without any bias or underlying relationship between them. It is assumed that both the samples are normally distributed, with similar variance and mean. In other words, it is assumed that both sets of observations follow a bell curve shape when plotted on a graph, with both having roughly equal spread around the mean value. Both samples should have equal sample sizes. This ensures that all values within both samples have an equal opportunity to contribute to the final result and no one value has greater influence than another. Conclusion In conclusion, a paired sample t-test is an effective statistical test used to compare the means of two samples. It should be used whenever you wish to determine if there is a statistically significant difference between two related or paired groups. The assumptions for this type of analysis must be met in order to get reliable and accurate results. This includes observations within the two samples being related or paired in a meaningful way, all observations being independent of each other, data normally distributed with similar variances and means, and equal sample sizes. With these assumptions in mind, a paired sample t-test can provide reliable and accurate results. I have been recently working in the area of Data analytics including Data Science and Machine Learning / Deep Learning. I am also passionate about different technologies including programming languages such as Java/JEE, Javascript, Python, R, Julia, etc, and technologies such as Blockchain, mobile computing, cloud-native technologies, application security, cloud computing platforms, big data, etc. I would love to connect with you on LinkedIn. Check out my latest book titled as First Principles Thinking: Building winning products using first principles thinking. The sample dataset has placement test scores (out of 100 points) for four subject areas: English, Reading, Math, and Writing. Students in the sample completed all 4 placement tests when they enrolled in the university. Suppose we are particularly interested in the English and Math sections, and want to determine whether students tended to score higher on their English or Math test, on average. We could use a paired t test to test if there was a significant difference in the average of the two tests. Before the Test Variable English has a high of 101.95 and a low of 59.83, while variable Math has a high of 93.78 and a low of 35.32 (Analyze > Descriptive Statistics > Descriptives). The mean English score is much higher than the mean Math score (82.79 versus 65.47). Additionally, there were 409 cases with non-missing English scores, and 422 cases with non-missing Math scores, but only 398 cases with non-missing observations for both variables. (Recall that the sample dataset has 435 cases in all.) Let's create a comparative boxplot of these variables to help visualize these numbers. Click Analyze > Descriptive Statistics > Explore. Add English and Math to the Dependents box; then, change the Display option to Plots. We'll also need to tell SPSS to put these two variables on the same chart. Click the Plots button, and in the Boxplots area, change the selection to Dependents Together. You can also uncheck Stem-and-leaf. Click Continue. Then click OK to run the procedure. We can see from the boxplot that the center of the English scores is much higher than the center of the Math scores, and that there is slightly more spread in the Math scores than in the English scores. Both variables appear to be symmetrically distributed. It's quite possible that the paired samples t test could come back significant. Running the Test Click Analyze > Compare Means and Proportions > Paired-Samples T Test. If you are using an older version of SPSS Statistics (prior to version 29), the menu path is Analyze > Compare Means > Paired-Samples T Test. Select the variable English and move it to the Variable1 slot in the Paired Variables box. Then select the variable Math and move it to the Variable2 slot in the Paired Variables box. Click OK. Syntax T-TEST PAIRS=English WITH Math (PAIRED) /CRITERIA=CI(9500) /MISSING=ANALYSIS. Output Tables There are three tables: Paired Samples Statistics, Paired Samples Correlations, and Paired Samples Test. Paired Samples Statistics gives univariate descriptive statistics (mean, sample size, standard deviation, and standard error) for each variable entered. Notice that the sample size here is 398; this is because the paired t-test can only use cases that have non-missing values for both variables. Paired Samples Correlations shows the bivariate Pearson correlation coefficient (with a two-tailed test of significance) for each pair of variables entered. Paired Samples Test gives the hypothesis test results. The Paired Samples Statistics output repeats what we examined before we ran the test. The Paired Samples Correlation table adds the information that English and Math scores are significantly positively correlated (r = .243). Why does SPSS report the correlation between the two variables when you run a Paired t Test? Although our primary interest when we run a Paired t Test is finding out if the means of the two variables are significantly different, it's also important to consider how strongly the two variables are associated with one another, especially when the variables being compared are pre-test/post-test measures. For more information about correlation, check out the Pearson Correlation tutorial. Reading from left to right: First column: The pair of variables being tested, and the order the subtraction was carried out. (If you have specified more than one variable pair, this table will have multiple rows.) Mean: The average difference between the two variables. Standard deviation: The standard deviation of the difference scores. Standard error mean: The standard error (standard deviation divided by the square root of the sample size). Used in computing both the test statistic and the upper and lower bounds of the confidence interval. t: The test statistic (denoted t) for the paired T test. df: The degrees of freedom for this test. Sig. (2-tailed): The p-value corresponding to the given test statistic t with degrees of freedom df. Decision and Conclusions From the results, we can say that: English and Math scores were weakly and positively correlated (r = 0.243, p < 0.001). There was a significant average difference between English and Math scores (t397 = 36.313, p < 0.001). On average, English scores were 17.3 points higher than Math scores (95% CI [16.36, 18.23]).