

PAIRED t-TEST:

A CASE STUDY: Employees performance data

13

Dr. D S Dhakre & Prof. D Bhattacharya
Visva Bharati, Sriniketan, West Bengal, India

A paired t-test, also known as a dependent t-test or paired samples t-test, is a statistical procedure used to determine if there is a significant difference between the means of two related groups. It is typically used when the observations within each group are paired or matched in some way.

Here's how the paired t-test works:

1.Data Structure: In a paired t-test, each subject or item in the sample contributes two measurements or scores, which are paired based on some characteristic (e.g., before and after treatment, measurements from the same subject at different times, etc.).

2.Null Hypothesis (H_0): The null hypothesis states that there is no significant difference between the means of the paired measurements. Mathematically, $H_0: \mu_d = 0$, where μ_d represents the mean of the differences between paired measurements.

3.Alternative Hypothesis (H_1): The alternative hypothesis proposes that there is a significant difference between the means of the paired measurements. Depending on the direction of interest, H_1 could be $\mu_d \neq 0$ (two-tailed test), $\mu_d > 0$ (one-tailed test for positive difference), or $\mu_d < 0$ (one-tailed test for negative difference).

4.Assumptions:

- The paired differences should be approximately normally distributed.
- The paired differences should be independent of each other.
- The underlying data from which the paired differences are derived should be interval or ratio scale.

5. Test Statistic: The test statistic for the paired t-test is calculated as:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

where:

- $\bar{d}$ is the mean of the paired differences.
- s_d is the standard deviation of the paired differences.
- n is the number of pairs.

6.Degrees of Freedom: The degrees of freedom df for the paired t-test is $n-1$, where n is the number of pairs.

7.Decision: Compare the calculated t-value to the critical t-value from the t-distribution table or use software to obtain the p-value. If the p-value is less than the significance level (often denoted as α , typically 0.05), then there is sufficient evidence to reject the null hypothesis and conclude that there is a significant difference between the means of the paired measurements.

Paired t-tests are commonly used in research and experimental studies where the same subjects are measured before and after an intervention, or under two different conditions. It allows researchers to assess whether the intervention or condition significantly affected the outcome of interest.

The Problem

The purpose of the study is to ascertain employee performance is improving or not.

Data

Download **Survey Data** file form the given link <https://dsdhakre.in/Datafiles.html>

Res	Dept	B_Train	A_Train	Tra
1	Production	7.5	8.0	Tra
2	Marketing	9.5	8.0	
3	Purchase & Store	9.5	10.0	
4	Production	9.5	7.5	
5	HR	6.5	6.0	
6	Production	7.0	9.0	
7	Purchase & Store	4.0	10.0	
8	Marketing	9.0	9.5	
9	Production	9.5	8.0	
10	Marketing	9.5	9.0	
11	Marketing	9.5	10.0	
12	Marketing	9.5	9.5	
13	Finance	9.5	7.5	
14	Marketing	9.5	9.0	
15	Marketing	9.5	7.5	
16	Production	9.0	9.5	
17	HR	9.5	8.0	
18	Finance	7.5	6.0	
19	Marketing	9.5	10.0	
20	Purchase & Store	2.0	8.0	
21	Purchase & Store	2.0	9.0	
22	HR	7.0	9.0	
23	Finance	7.5	8.0	
24	Finance	9.5	9.5	
25	Marketing	9.0	10.0	
26	Finance	7.5	9.0	

Hypothesis

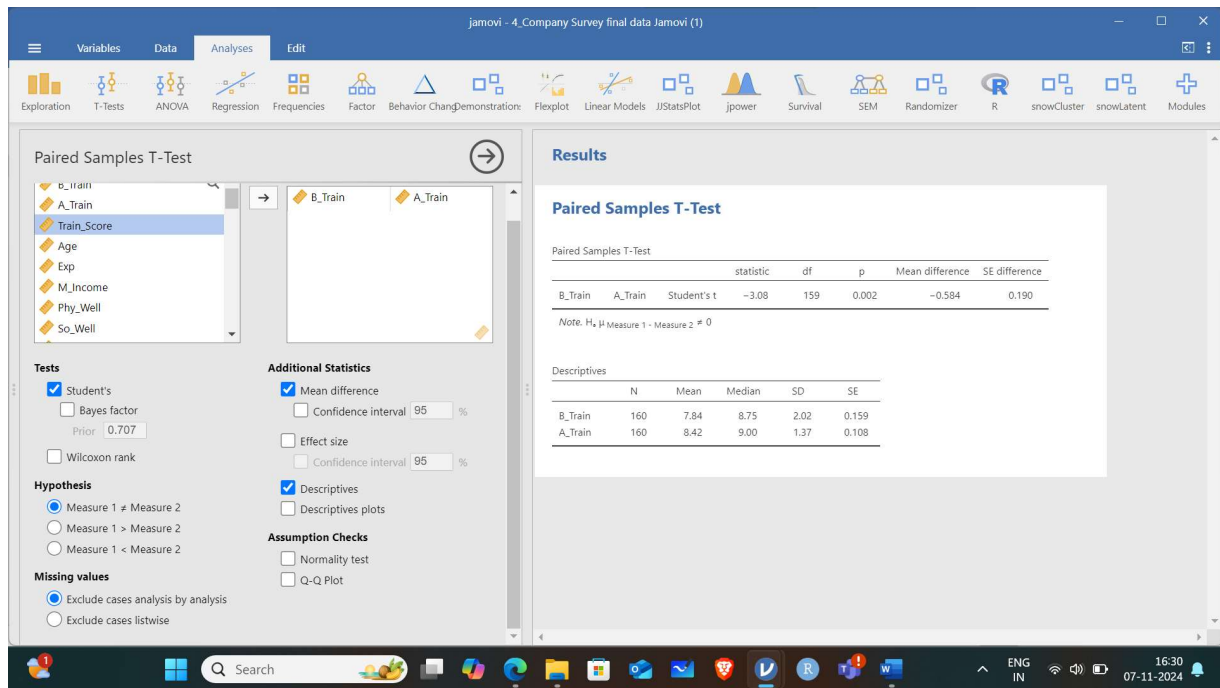
H_0 : Performance of employees is same in both test

Independent Variable: **Test**

Dependent variable: **Marks**

Step

The screenshot displays the Jamovi software interface. The main window shows a dataset with columns: 'train', 'A_Train', and 'Train_Sc'. The 'train' column contains numerical values (7.5, 9.5, 9.5, 9.5, 6.5, 7.0, 4.0, 9.0, 9.5, 9.5, 9.5, 9.5, 9.5, 9.5, 9.0, 9.5, 7.5, 9.5, 9.5, 2.0, 2.0, 7.0, 7.5, 9.5, 9.0, 7.5). The 'A_Train' column contains numerical values (8.0, 8.0, 10.0, 7.5, 6.0, 9.0, 10.0, 9.5, 8.0, 9.0, 10.0, 9.5, 7.5, 9.0, 9.5, 6.0, 10.0, 8.0, 9.0, 9.5, 9.0, 10.0, 8.0, 8.0, 9.0, 8.0, 9.5, 10.0, 9.0). The 'Train_Sc' column contains numerical values (8.0, 8.0, 10.0, 7.5, 6.0, 9.0, 10.0, 9.5, 8.0, 9.0, 10.0, 9.5, 7.5, 9.0, 9.5, 6.0, 10.0, 8.0, 9.0, 9.5, 9.0, 10.0, 8.0, 9.0, 9.0, 9.5, 10.0, 9.0). A dropdown menu is open under the 'Analyses' tab, showing options: 'Independent Samples T-Test', 'Paired Samples T-Test', 'One Sample T-Test', 'Bayesian (post)', 'Bayesian Independent Samples T-Test', 'Bayesian Paired Samples T-Test', and 'Bayesian One Sample T-Test'. The 'Results' panel on the right is empty. The bottom status bar shows 'Row count: 160', 'Filtered: 0', 'Deleted: 0', 'Added: 0', and 'Cells edited: 38'. The Windows taskbar at the bottom shows the date '07-11-2024' and time '16:27'.



Reporting

Group	Mean	Sd	df	SEd	t-value	sig
Before training Marks	7.84	2.02	159	0.190	-3.08	0.002
After training Marks	8.42	1.37				

So, the p-value found significant at 1% level of significance, Hence Null hypothesis H_0 is rejected. Here, the mean scores show that after training, the scores were increased. It means performance of employee is not same in both the tests. Hence **Employees are improving their knowledge.**

Do your self

Download **paired t test** file form the given link <https://dshakre.in/Datafiles.html>