Function of the test:

The Paired Samples *t*-test compares two means that are from the same individual, object, or related units. The two means typically represent two different times (e.g., pre-test and post-test with an intervention between the two time points) or two different but related conditions or units (e.g., left and right ears, twins).

The purpose of the test is to determine whether there is statistical evidence that the mean difference between paired observations on a particular outcome is significantly different from zero. The Paired Samples *t* Test is a parametric test.

Test Assumptions

1. Dependent variable is continuous (interval or ratio scale data)
2. Sample must be related or dependent. This means that the subjects in the first group are also in the second group.
3. Random sample of data
4. Normal distribution (approximately) of the difference between the paired values
5. No outliers in the difference between the two related groups

Null and Alternative Hypothesis:

Let D be the difference (Before – After) between two sets of observations in the population. If X_1 = measurements in first occasion and X_2 = measurements in second occasion, then D = $X_1 - X_2$. Let μ_D be the average difference of two sets of measurements i.e. average of the difference D.

Then null and alternative hypothesis are written as follows:

Null Hypothesis	**Alternative Hypothesis**	**No. of tails**
$H_0: \mu_D = 0$ Or H_0: $\mu_1 - \mu_2 = 0$ ("the difference between the paired population means is equal to 0") H_1: $\mu_1 - \mu_2 \neq 0$ ("the difference between the paired population means is not 0")	$H_1: \mu_D \neq 0$	Two tailed test
$H_0: \mu_D \geq 0$	$H_1: \mu_D < 0$	Lower tailed test
$H_0: \mu_D \leq 0$	$H_1: \mu_D > 0$	Right tailed test

Test Statistics

The test statistics for the paired sample t-test is given by,

$$t = \frac{\bar{x}_d - \mu_D}{s_d / \sqrt{n}} = \frac{\bar{x}_d}{s_d / \sqrt{n}}$$

Under null hypothesis H_0, the value of population mean of difference μ_D is assumed to be zero.

where

$\bar{d}$ or $\bar{x}_d$ = Sample mean of the differences
n = Sample size (i.e., number of observations)
s_d = Sample standard deviation of the differences

$s_d / \sqrt{n}$ = Standard error of the sample differences d

The test statistics follows Student's t distribution with n – 1 degrees of freedom.

Decision Rule:

We adopt following decision rule:

Hypothesis	
Case I (Two-sided test)	Reject H_0 if $\lvert$cal t$\rvert \geq t_{\alpha/2}(n-1)$
Case II (Left sided test)	Reject H_0 if cal t $\leq - t_{\alpha}(n-1)$
Case III (Right sided test)	Reject H_0 if cal t $\geq + t_{\alpha}(n-1)$

<u>Paired t – test example</u>

Fifteen secretaries of a certain corporation were sent for a two-day training to increase their typing skills. The table below shows the typing speed of the secretaries in words per minute before and after the training?

Secretary	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Before	75	60	54	67	60	85	60	74	69	85	82	68	70	58	72
After	75	65	59	66	65	86	70	71	68	86	80	70	75	57	80

Assuming the data to be normally distributed and α = 0.05, is there an evidence for the training improved the typing speed of the secretaries. Use paired sample t test.

Solution

Let X_1 be typing speed of the secretaries (in minutes) before training and X_2 be the typing speed after training

Paired sample size (n) = 15

Step 1: Null and Alternative Hypothesis

Let D be the difference (Before – After) between two sets of observations in the population. If X_1 = measurements in first occasion and X_2 = measurements in second occasion, then $D = X_1 - X_2$. Let μ_D be the average difference of two sets of measurements i.e. average of the difference D.

H_0: $\mu_D \geq 0$ (Two-day training is not effective)

H_1 : $\mu_D < 0$ (Two-day training is effective)

Step2: Choice of α for the test

The level of significance for the test is 5 % (Let)

Step 3: Test statistic

The appropriated test statistics for the paired sample t-test is given by,

$$t = \frac{\overline{x}_d - \mu_D}{s_d / \sqrt{n}} = \frac{\overline{x}_d}{s_d / \sqrt{n}}$$

Under null hypothesis H_0 , the value of population mean of difference μ_D is assumed to be zero.

where,

$\overline{d}$ or $\overline{x}_d$ = Sample mean of the differences
n = Sample size (i.e., number of observations)
s_d = Sample standard deviation of the differences

$s_d / \sqrt{n}$ = Standard error of the sample differences d

The test statistics follows Student's t distribution with n – 1 degrees of freedom.

Step 4: Tabulated t or Critical t

The test is left sided and level of significance chosen for the test is 5 %

The rejection region lies in the left side of the t curve having n – 1 degree of freed.

Deegrees of freedom = n – 1 = 15 -1 =14

Critical value from the t table is given by

tc = $t_{0.05}(14)$= 1.761

AR : cal t ≥ - 1.761

RR: cal t < - 1.761

Step 5: Calculated t or observed t

Before (X1)	After (X2)	diff (d)	d^2
75	75	0	0
60	65	-5	25
54	59	-5	25
67	66	1	1
60	65	-5	25
85	86	-1	1
60	70	-10	100
74	71	3	9
69	68	1	1
85	86	-1	1
82	80	2	4
68	70	-2	4
70	75	-5	25
58	57	1	1
72	80	-8	64
		-34	286

$$\bar{x}_d = \frac{\sum d}{n} = -34/15 = -2.2667$$

$$s_d = \sqrt{\frac{1}{n-1}\left\{\sum d^2 - n\ .\ \bar{d}^2\right\}} = 3.8631$$

Now,

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} = \frac{-2.2667}{3.8631/\sqrt{15}} = -2.2725$$

Hence cal t = - 2.2725

Step 6: Statistical Decision

Since cal t = - 2.2725 falls in the acceptance region (t ≥ - 1.761) we do not reject null hypothesis at 5 % level of significance

conclusion: The staffs are not benefited by two-day training on improving the typing speed.

Problem:

To test a theory that alcohol consumption can have an effect on test score, a researcher conducts a study on ten adults. Each is given a test. Then for one week, each subject is required to consume a certain amount of alcohol, then he/she is retested. The results are shown below.

Subject	1	2	3	4	5	6	7	8	9	10
Score before	105	109	98	112	109	117	123	114	95	101
Score after	106	105	94	109	105	115	125	114	98	100

At $\alpha = 0.05$, test the claim that alcohol does not affect a person's test scores.