Biostatistics

t-test
Inferences about Population Means

:Deepa Anwar

Goals :

The basic goal of these simple tests is to show that the distribution of the given data under examination are not produced by chance and that there is some systematic pattern therein.

Main point is to show the mean of a sample is reflective of the population

A **t-test** is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. The **t-test** is one of many **tests** used for the purpose of hypothesis **testing** in statistics.

- The *t*-test is used to test hypotheses about means when the population variance is unknown (the usual case). Closely related to z, the unit normal.
- Developed by Gossett for the quality control of beer.
- Comes in 3 varieties:
  Single sample, independent samples, and dependent samples.

What kind of *t* is it?

- Single sample *t* – we have only 1 group; want to test against a hypothetical mean.

- Independent samples *t* – we have 2 means, 2 groups; no relation between groups, e.g., people randomly assigned to a single group.

- Dependent *t* – we have two means. Either same people in both groups, or people are related, e.g., husband-wife, left hand-right hand, hospital patient and visitor.

Paired Samples t-Test:

Solved example:

Step 1:

Subject #	Score 1	Score 2
1	3	20
2	3	13
3	3	13
4	12	20
5	15	29
6	16	32
7	17	23
8	19	20
9	23	25
10	24	15
11	32	30

step 2:

Subject #	Score 1	Score 2	X-Y
1	3	20	-17
2	3	13	-10
3	3	13	-10
4	12	20	-8
5	15	29	-14
6	16	32	-16
7	17	23	-6
8	19	20	-1
9	23	25	-2
10	24	15	9
11	32	30	2
		SUM:	-73

step 3

Subject #	Score 1	Score 2	X-Y	(X-Y)^2
1	3	20	-17	289
2	3	13	-10	100
3	3	13	-10	100
4	12	20	-8	64
5	15	29	-14	196
6	16	32	-16	256
7	17	23	-6	36
8	19	20	-1	1
9	23	25	-2	4
10	24	15	9	81
11	32	30	2	4
		SUM:	-73	1131

Step 4

$$t = \frac{(\Sigma D)/N}{\sqrt{\frac{\Sigma D^2 - \frac{((\Sigma D)^2)}{N}}{(N-1)(N)}}}$$

ΣD: Sum of the differences (Sum of X-Y from Step 2)
ΣD^2: Sum of the squared differences (from Step 3)
$(\Sigma D)^2$: Sum of the differences (from Step 2), squared.

Step 5:

$$t = \cfrac{(\Sigma D)/N}{\sqrt{\cfrac{\Sigma D^2 - \left(\cfrac{(\Sigma D)^2}{N}\right)}{(N-1)(N)}}}$$

$$t = \cfrac{-73/11}{\sqrt{\cfrac{1131 - \left(\cfrac{(-73)^2}{11}\right)}{(11-1)(11)}}}$$

$$t = \cfrac{-73/11}{\sqrt{\cfrac{1131 - \left(\cfrac{5329}{11}\right)}{110}}}$$

$$t = -\ 2.74$$

Step 6: Subtract 1 from the sample size to get the degrees of freedom. We have 11 items, so 11-1 = 10.

Step 7: Find the p-value in the t-table, using the degrees of freedom in Step 6. If you don't have a specified alpha level, use 0.05 (5%). For this sample problem, with df=10, the t-value is 2.228.

Step 8: Compare your t-table value from Step 7 (2.228) to your calculated t-value (-2.74). The calculated t-value is greater than the table value at an alpha level of .05. The p-value is less than the alpha level: p <.05. We can reject the null hypothesis that there is no difference between means.

Note: You can ignore the minus sign when comparing the two t-values, as ± indicates the direction; the p-value remains the same for both directions.