STUDENT'S T TEST

Definition

- The **t-test** (also called the **Student t-test**) is **used to** determine the significance of the difference between the means of two sets of data A and B
- The question asked is: Is the mean of sample A significantly different from the mean of sample B?
- The **t**-statistic was introduced in 1908 by William Sealy Gosset, a chemist
- t- test depends on the properties of Normal distribution curves
- The null hypothesis (H_0) states that the means of samples A and B are the same.
- Alternative hypothesis (H_A): means of samples A and B are different

Computation of t-test

- Formula used is

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{(s^2(\frac{1}{n_1} + \frac{1}{n_2}))}}$$

- t is the t-value,
- $\overline{x}_1$ and $\overline{x}_2$ are the means of the two groups being compared,
- s^2 is the common variance of the two groups, and
- n_1 and n_2 are the number of observations in each of the groups.

Example-comparison of wing length

Sample A (wing length of males in mm)	Sample B (wing length of females in mm)
73.0	71.0
74.2	71.5
75.0	72.0
75.3	72.4
75.5	73.5
75.8	74.0
	74.3
	75.2
$n_1 = 6$	$n_2 = 8$
$\bar{X}_1 = 74.8$	$\bar{X}_2 = 72.99$
$s_1 = 1.04$	$s_2 = 1.48$
$s_1^2 = 1.08$	$s_2^2 = 2.20$

$s^2 = (6X1.08)+(8X2.20)/6+8-2$

$=6.48 +17.6/12$

$=24.1/12$

$=2.01$

$s=\sqrt{2.01} = 1.42$

Calculation of 't' on next slide

Example-comparison of wing length

- Substituting the value of 's' and $\bar{x}$ in the formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s^2(\frac{1}{n_1} + \frac{1}{n_2}))}}$$

- t= 74.8 – 72.99/1.42√1/6 +1/8
- = 1.81/0.767
- t = 2.36 with 6+8-2 degrees of freedom
- The calculated value of 't' is greater than the tabulated value of 2.179 at P=0.05 for 12 degrees of freedom
- Therefore the null hypothesis (H_o) is rejected
- Conclusion: the difference between the means of samples A and B is statistically significant

t Distribution table

df/α	0.05	0.02	0.01	0.001
1	12.706	31.821	64	637
2	4.303	6.965	10	31.598
3	3.182	4.541	5.841	12.929
4	2.776	3.747	4.604	8.61
5	2.571	3.365	4.032	6.869
6	2.447	3.143	3.707	5.959
7	2.365	2.998	3.499	5.408
8	2.306	2.896	3.355	5.041
9	2.263	2.821	3.25	4.781
10	2.228	2.764	3.169	4.587
11	2.201	2.718	3.106	4.437
12	2.179	2.681	3.055	4.318
13	2.16	2.65	3.012	4.221
14	2.145	2.624	2.977	4.14

The calculated value of 't' (= 2.36) is greater than the tabulated value of 2.179 at P=0.05 for 12 degrees of freedom

One-tailed test

- In one-tailed test, the alternative hypothesis (H_A) is that the mean of a particular nominated sample (A or B) will be greater than the mean of the other sample (B or A).
- The critical value of 't' is lower in one-tailed test

Two-tailed test

- In two-tailed test, the alternative hypothesis (H_A) is that means of the two samples A and B are merely different.
- Two-tailed test is more stringent and thus recommended.

Restrictions and cautions

- The t-test assumes that data are measured at interval/ratio level.
- Data should be derived from Normally distributed populations.
- Counts data may not be Normally distributed, so count data should be transformed logarithmically before performing t-test.
- Proportions and percentages data need to be arcsine transformed.