## * t-test for testing the significance of difference between two means:

(OR testing the significance of diff betⁿ mean of two independent samples)

Given two independent random samples of sizes $n_1$ and $n_2$ with sample means $\bar{X}_1$ and $\bar{X}_2$ and sample standard deviations $S_1$ and $S_2$, we may be interested in testing the hypothesis that both samples have come from the same normal population.

i.e We may want to test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$

where $\mu_1$ and $\mu_2$ are the population means.

We use the following test statistic.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where. $\bar{X}_1$ = mean of the first sample

$\bar{X}_2$ = mean of the second sample

$n_1$ = number of observations in the first sample.

$n_2$ = number of observations in the second sample.

$S$ = combined standard deviation.

The value of $S$ can be calculated by

$$S = \sqrt{\frac{\Sigma(X_1 - \bar{X}_1)^2 + \Sigma(X_2 - \bar{X}_2)^2}{n_1 + n_2 - 2}}$$

$$\therefore\ S = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$$

where $S_1$ and $S_2$ are S.D of first and second sample respectively.

OR

$$S = \sqrt{\frac{1}{n_1+n_2-2}\left\{\Sigma d_1^2 - \frac{(\Sigma d_1)^2}{n_1} + \Sigma d_2^2 - \frac{(\Sigma d_2)^2}{n_2}\right\}}$$

where $d_1 = X_1 - A$ and $d_2 = X_2 - B$, A and B being the assumed means for the first and second sample.

The degrees of freedom $= n_1 + n_2 - 2$

From the table of t-distribution, we find the value $t_{n_1+n_2-2,\ 0.05}$. If $|t| \le t_{n_1+n_2-2,\ 0.05}$;

we accept $H_0$ otherwise we reject it at 5% level of significance.

Ex-1 Two types of drugs were used on 5 and 7 patients for reducing their weights. Drug A is imported and drug B indigeneous. The decrease in the weight after using the drugs for six months was recorded as given below.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Drug A | 11 | 13 | 12 | 14 | 10 | | |
| Drug B. | 12 | 9 | 8 | 15 | 14 | 9 | 10 |

Is there significant difference in the efficacy of two drugs? If not, which drug should you buy?

Sol$^n$: Let $H_0$: There is no significant diff in the efficacy of the two drugs.

$H_1$: There is significant diff in the efficacy of two drugs.

Now

| $x_1$ | $x_1 - \bar{x}_1$ | $(x_1-\bar{x}_1)^2$ | $x_2$ | $x_2 - \bar{x}_2$ | $(x_2-\bar{x}_2)^2$ |
|---|---|---|---|---|---|
| 11 | -1 | 1 | 12 | 1 | 1 |
| 13 | 1 | 1 | 9 | -2 | 4 |
| 12 | 0 | 0 | 8 | -3 | 9 |
| 14 | 2 | 4 | 15 | 4 | 16 |
| 10 | -2 | 4 | 14 | 3 | 9 |
| | | | 9 | -2 | 4 |
| | | | 10 | -1 | 1 |
| $\Sigma x_i = 60$ | | $\Sigma(x_1-\bar{x}_1)^2 = 10$ | $\Sigma x_2 = 77$ | | $\Sigma(x_2-\bar{x}_2)^2 = 44$ |

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{60}{5} = 12, \quad \bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{77}{7} = 11$$

$$S = \sqrt{\frac{\Sigma(x_1-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{10+44}{5+7-2}} = \sqrt{\frac{54}{10}} = 2.3237$$

$$\therefore |t| = \left|\frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right| = \left|\frac{12-11}{2.3237\sqrt{\frac{1}{5}+\frac{1}{7}}}\right| = 0.735$$

$$= 0.7350$$

$d.f = n_1 + n_2 - 2 = 5 + 7 - 2 = 10$.

From the table of t-distribution we obtain

$t_{10, 0.05} = 2.225$.

Here $|t| < t_{10, 0.05}$

$\therefore$ we accept $H_0$ at 5% level of significance. and conclude that there so no significant difference in the efficacy of two drugs.

**Ex-2** Blood glucose level of pigeons is compared with rabbits. Apply proper statistical test to know the significance of difference of blood glucose levels of the two using the following data and comment on your result.

| Sr. No | Blood glucose level per 100 ml | |
|---|---|---|
| | pigeons | Rabbits. |
| 1 | 200 | 145 |
| 2 | 186 | 125 |
| 3 | 176 | 100 |
| 4 | 184 | 112 |
| 5 | 170 | 127 |
| 6 | 172 | 139 |
| 7 | 170 | 151 |
| 8 | 163 | 140 |
| 9 | 176 | 159 |
| 10 | 173. | 132. |

$Sol^n$: Let $x_1$ = Blood glucose level of pigeons.
$x_2$ = Blood glucose level of rabbits.

| $x_1$ | $x_1-\bar{x}_1$ | $(x_1-\bar{x}_1)^2$ | $x_2$ | $x_2-\bar{x}_2$ | $(x_2-\bar{x}_2)^2$ |
|---|---|---|---|---|---|
| 200 | 23 | 529 | 145 | 12 | 144 |
| 186 | 9 | 81 | 125 | -8 | 64 |
| 176 | -1 | 1 | 100 | -33 | 1089 |
| 184 | 7 | 49 | 112 | -21 | 441 |
| 170 | -7 | 49 | 127 | -6 | 36 |
| 172 | -5 | 25 | 139 | 6 | 36 |
| 170 | -7 | 49 | 151 | 18 | 324 |
| 163 | -14 | 196 | 140 | 7 | 49 |
| 176 | -1 | 1 | 159 | 26 | 676 |
| 173 | -4 | 16 | 132 | -1 | 1 |
| $\Sigma x_1 = 1770$ | | $\Sigma(x_1-\bar{x}_1)^2 = 996$ | $\Sigma x_2 = 1330$ | | $\Sigma(x_2-\bar{x}_2)^2 = 2860$ |

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{1770}{10} = 177, \quad \bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{1330}{10} = 133$$

$$s = \sqrt{\frac{\Sigma(x_1-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{996+2860}{10+10-2}} = \sqrt{\frac{3856}{18}} = 14.6363.$$

Let $H_0$: There is no significant difference in blood glucose levels of pigeons and rabbits.
$H_1$: There is significant difference in blood glucose levels of pigeons and rabbits.

$$|t| = \left| \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| = \left| \frac{177 - 133}{14.6363\sqrt{\frac{1}{10} + \frac{1}{10}}} \right|$$

$$= \left| \frac{44 \times \sqrt{5}}{14.6363} \right|$$

$$= 6.72$$

$d.f = n_1 + n_2 - 2 = 10 + 10 - 2 = 18$

From the table of t-distribution, we find.

$t_{18,\,0.05} = 2.101$

$|t| > t_{18,\,0.05}$

$\therefore$ We reject $H_0$ at 5% level of significance and conclude that there is significant difference in blood glucose levels of pigeons and rabbits.

Ex-3. A large group of teachers are trained under QIP (Quality Improvement programme) where some are trained by Institution A and some are trained by institution B. In a random sample of 10 teachers taken from a large group, the following marks are obtained in an appropriate achivement test.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Institute A | 65 | 69 | 73 | 71 | 75 | 66 | 71 | 68 | 68 | 74 |
| Institute B: | 78 | 69 | 72 | 77 | 84 | 70 | 73 | 77 | 75 | 65 |

Test the claim that institution B is more effective at 0.05 level of significance under the assumption that the two populations are normally distributed with same variances.

Sol$^n$

| $x_A$ | $x_A - \bar{x}_A$ | $(x_A - \bar{x}_A)^2$ | $x_B$ | $x_B - \bar{x}_B$ | $(x_B - \bar{x}_B)^2$ |
|---|---|---|---|---|---|
| 65 | -5 | 25 | 78 | 4 | 16 |
| 69 | -1 | 1 | 69 | -5 | 25 |
| 73 | 3 | 9 | 72 | -2 | 4 |
| 71 | 1 | 1 | 77 | 3 | 9 |
| 75 | 5 | 25 | 84 | 10 | 100 |
| 66 | -4 | 16 | 70 | -4 | 16 |
| 71 | 1 | 1 | 73 | -1 | 1 |
| 68 | -2 | 4 | 77 | 3 | 9 |
| 68 | -2 | 4 | 75 | 1 | 1 |
| 74 | 4 | 16 | 65 | -9 | 81 |
| $\Sigma x_A = 700$ | | $\Sigma(x_A - \bar{x}_A)^2 = 102$ | $\Sigma x_B = 740$ | | $\Sigma(x_B - \bar{x}_B)^2 = 262$ |

Let $x_A$ = marks obtained in appropriate achievement test by teachers trained by institute A.

$x_B$ = " " " by institute B.

$n_A = n_B = 10$.

$\bar{x}_A = \frac{\Sigma x_A}{n_A} = \frac{700}{10} = 70$, $\bar{x}_B = \frac{\Sigma x_B}{n_B} = \frac{740}{10} = 74$

$$S = \sqrt{\frac{\Sigma(x_A - \bar{x}_A)^2 + \Sigma(x_B - \bar{x}_B)^2}{n_A + n_B - 2}}$$

$$= \sqrt{\frac{102 + 262}{10 + 10 - 2}} = \sqrt{\frac{364}{18}} = \sqrt{20.22} = 4.50$$

Null hypothesis $H_0 : \mu_A = \mu_B$ (there is no difference in teaching by institute A and B)

Alternative hypothesis $H_1 : \mu_A < \mu_B$ (Institute B is more effective than institute A)
(Left one tailed test.)

Now

$$t = \frac{\bar{x}_A - \bar{x}_B}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{70 - 74}{(4.50)\sqrt{\frac{1}{10} + \frac{1}{10}}} = \frac{-4}{2.012} = -1.988$$

Here $df = n_A + n_B - 2 = 10 + 10 - 2 = 18$

From the table of t-distribution we find

$t_{18, 0.05} = -1.734$

$\therefore t < t_{18, 0.05}$ (left side or negative side)

$\therefore$ We reject $H_0$ at 5% level of significance.
Thus, we conclude that the institution B is more effective than the institution A in teaching.

Note: In the Ex-2 & Ex-3 there are equal number of observations in the two samples, we cannot use the paired t-test in this two examples, because the observations in the two samples are not paired. i.e they are not on the same units.