**Expected value:**

$$t_e = \left|\frac{\bar{x}-\mu}{s/\sqrt{n-1}}\right| \quad \text{follows t-distribution with } (26-1) = 25\text{d.f}$$

$$= 1.708$$

**Inference:**

Since $t_0 > t_e$, $H_0$ is rejected at 5% level of significance. Hence we conclude that advertisement is certainly effective in increasing the sales.

### 6.3 Test of significance for difference between two means:

#### 6.3.1 Independent samples:

Suppose we want to test if two independent samples have been drawn from two normal populations having the same means, the population variances being equal. Let $x_1, x_2, \ldots x_{n_1}$ and $y_1, y_2, \ldots y_{n_2}$ be two independent random samples from the given normal populations.

**Null hypothesis:**

$H_0 : \mu_1 = \mu_2$ i.e. the samples have been drawn from the normal populations with same means.

**Alternative Hypothesis:**

$H_1 : \mu_1 \neq \mu_2$ $(\mu_1 < \mu_2$ or $\mu_1 > \mu_2)$

**Test statistic:**

Under the $H_0$, the test statistic is

$$t_0 = \left|\frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|$$

$$\text{where } \bar{x} = \frac{\Sigma x}{n_1} ;\ \bar{y} = \frac{\Sigma y}{n_2}$$

$$\text{and } S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right] = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}$$

**Expected value:**

$$t_e = \left| \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \quad \text{follows t-distribution with } n_1 + n_2 - 2 \text{ d.f}$$

**Inference:**

If the $t_0 < t_e$ we accept the null hypothesis. If $t_0 > t_e$ we reject the null hypothesis.

**Example 3:**

A group of 5 patients treated with medicine 'A' weigh 42, 39, 48, 60 and 41 kgs: Second group of 7 patients from the same hospital treated with medicine 'B' weigh 38, 42 , 56, 64, 68, 69 and 62 kgs. Do you agree with the claim that medicine 'B' increases the weight significantly?

**Solution:**

Let the weights (in kgs) of the patients treated with medicines A and B be denoted by variables X and Y respectively.

**Null hypothesis:**

$H_0 : \mu_1 = \mu_2$

i.e. There is no significant difference between the medicines A and B as regards their effect on increase in weight.

**Alternative Hypothesis:**

$H_1 : \mu_1 < \mu_2$ (left-tail) i.e. medicine B increases the weight significantly.

**Level of significance :** Let $\alpha = 0.05$

**Computation of sample means and S.Ds**

**Medicine A**

| X | $x - \bar{x}$ $(\bar{x} = 46)$ | $(x - \bar{x})^2$ |
|---|---|---|
| 42 | $-4$ | 16 |
| 39 | $-7$ | 49 |
| 48 | 2 | 4 |
| 60 | 14 | 196 |
| 41 | $-5$ | 25 |
| **230** | **0** | **290** |

$$\bar{x} = \frac{\Sigma x}{n_1} = \frac{230}{5} = 46$$

**Medicine B**

| Y | $y-\bar{y}$ $(\bar{y}=57)$ | $(y-\bar{y})^2$ |
|---|---|---|
| 38 | –19 | 361 |
| 42 | –15 | 225 |
| 56 | –1 | 1 |
| 64 | 7 | 49 |
| 68 | 11 | 121 |
| 69 | 12 | 144 |
| 62 | 5 | 25 |
| **399** | **0** | **926** |

$$\bar{y} = \frac{\Sigma y}{n_2} = \frac{399}{7} = 57$$

$$S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right]$$

$$= \frac{1}{10}\left[290 + 926\right] = 121.6$$

**Calculation of statistic:**

Under $H_0$ the test statistic is

$$t_0 = \left| \frac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \right|$$

$$= \left| \frac{46-57}{\sqrt{121.6\left(\frac{1}{5}+\frac{1}{7}\right)}} \right|$$

$$= \frac{11}{\sqrt{121.6 \times \frac{12}{35}}}$$

$$= \frac{11}{6.57} = 1.7$$

**Expected value:**

$$t_e = \left| \frac{\bar{x} - \bar{y}}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right|$$ follows t-distribution with (5+7–2) =10 d.f

$= 1.812$

**Inference:**

Since $t_0 < t_e$ it is not significant. Hence $H_0$ is accepted and we conclude that the medicines A and B do not differ significantly as regards their effect on increase in weight.

**Example 4:**

Two types of batteries are tested for their length of life and the following data are obtained:

| | No of samples | Mean life (in hrs) | Variance |
|---|---|---|---|
| Type A | 9 | 600 | 121 |
| Type B | 8 | 640 | 144 |

Is there a significant difference in the two means?

**Solution:**

We are given

$n_1=9$; $\bar{x}_1=600$hrs; $s_1^2=121$; $n_2=8$; $\bar{x}_2=640$hrs; $s_2^2=144$

**Null hypothesis:**

$H_0 : \mu_1 = \mu_2$ i.e. Two types of batteries A and B are identical i.e. there is no significant difference between two types of batteries.

**Inference:**

Since $t_0 > t_e$ it is highly significant. Hence $H_0$ is rejected and we conclude that the two types of batteries differ significantly as regards their length of life.

### 6.3.2 Related samples –Paired t-test:

In the t-test for difference of means, the two samples were independent of each other. Let us now take a particular situations where

(i) The sample sizes are equal; i.e., $n_1 = n_2 = n$(say), and

(ii) The sample observations ($x_1$, $x_2$, ……$x_n$) and ($y_1$, $y_2$, …..$y_n$) are not completely independent but they are dependent in pairs.

That is we are making two observations one before treatment and another after the treatment on the same individual. For example a business concern wants to find if a particular media of promoting sales of a product, say door to door canvassing or advertisement in papers or through T.V. is really effective. Similarly a pharmaceutical company wants to test the efficiency of a particular drug, say for inducing sleep after the drug is given. For testing of such claims gives rise to situations in (i) and (ii) above, we apply paired t-test.

**Paired – t –test:**

Let $d_i = X_i - Y_i$ ($i$ = 1, 2, ….n) denote the difference in the observations for the $i^{th}$ unit.

**Null hypothesis:**

$H_0 : \mu_1 = \mu_2$ ie the increments are just by chance

**Alternative Hypothesis:**

$H_1 : \mu_1 \neq \mu_2$ ( $\mu_1 > \mu_2$ (or) $\mu_1 < \mu_2$)

**Calculation of test statistic:**

$$t_0 = \left|\frac{\bar{d}}{S/\sqrt{n}}\right|$$

where $\bar{d} = \frac{\Sigma d}{n}$ and $S^2 = \frac{1}{n-1}\Sigma(d-\bar{d})^2 = \frac{1}{n-1}[\Sigma d^2 - \frac{(\Sigma d)^2}{n}]$

**Expected value:**

$$t_e = \left| \frac{\bar{d}}{S/\sqrt{n}} \right| \text{ follows t-distribution with } n-1 \text{ d.f}$$

**Inference:**

By comparing $t_0$ and $t_e$ at the desired level of significance, usually 5% or 1%, we reject or accept the null hypothesis.

**Example 5:**

To test the desirability of a certain modification in typists desks, 9 typists were given two tests of as nearly as possible the same nature, one on the desk in use and the other on the new type. The following difference in the number of words typed per minute were recorded:

| Typists | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| Increase in number of words | 2 | 4 | 0 | 3 | –1 | 4 | –3 | 2 | 5 |

Do the data indicate the modification in desk promotes speed in typing?

**Solution:**

**Null hypothesis:**

$H_0 : \mu_1 = \mu_2$ i.e. the modification in desk does not promote speed in typing.

**Alternative Hypothesis:**

$H_1 : \mu_1 < \mu_2$ (Left tailed test)

**Level of significance:** Let $\alpha = 0.05$

| Typist | d | $d^2$ |
|---|---|---|
| A | 2 | 4 |
| B | 4 | 16 |
| C | 0 | 0 |
| D | 3 | 9 |
| E | –1 | 1 |
| F | 4 | 16 |
| G | –3 | 9 |
| H | 2 | 4 |
| I | 5 | 25 |
| | $\Sigma d = 16$ | $\Sigma d^2 = 84$ |

$$\bar{d} = \frac{\sum d}{n} = \frac{16}{9} = 1.778$$

$$S = \sqrt{\frac{1}{n-1}[\sum d^2 - \frac{(\sum d)^2}{n}]}$$

$$= \sqrt{\frac{1}{8}[84 - \frac{(16)^2}{9}]} = \sqrt{6.9} = 2.635$$

**Calculation of statistic:**

Under $H_0$ the test statistic is

$$t_0 = \left|\frac{\bar{d}.\sqrt{n}}{S}\right| = \frac{1.778 \times 3}{2.635} = 2.024$$

**Expected value:**

$t_e = \left|\frac{\bar{d}.\sqrt{n}}{S}\right|$ follows t- distribution with $9 - 1 = 8$ d.f

$= 1.860$

**Inference:**

When $t_0 < t_e$ the null hypothesis is accepted. The data does not indicate that the modification in desk promotes speed in typing.

**Example 6:**

An IQ test was administered to 5 persons before and after they were trained. The results are given below:

| Candidates | I | II | III | IV | V |
|---|---|---|---|---|---|
| IQ before training | 110 | 120 | 123 | 132 | 125 |
| IQ after training | 120 | 118 | 125 | 136 | 121 |

Test whether there is any change in IQ after the training programme (test at 1% level of significance)

**Solution:**

**Null hypothesis:**

$H_0 : \mu_1 = \mu_2$ i.e. there is no significant change in IQ after the training programme.

**Alternative Hypothesis:**

$H_1 : \mu_1 \neq \mu_2$ (two tailed test)

**Level of significance :**

$\alpha = 0.01$

| | | | | | | |
|---|---|---|---|---|---|---|
| x | 110 | 120 | 123 | 132 | 125 | Total |
| y | 120 | 118 | 125 | 136 | 121 | - |
| d = x–y | –10 | 2 | –2 | –4 | 4 | –10 |
| $d^2$ | 100 | 4 | 4 | 16 | 16 | 140 |

$$\bar{d} = \frac{\Sigma d}{n} = \frac{-10}{5} = -2$$

$$S^2 = \frac{1}{n-1}\left[\Sigma d^2 - \frac{(\Sigma d)^2}{n}\right]$$

$$= \frac{1}{4}\left[140 - \frac{100}{5}\right] = 30$$

**Calculation of Statistic:**

Under $H_0$ the test statistic is

$$t_0 = \left|\frac{\bar{d}}{S/\sqrt{n}}\right|$$

$$= \left|\frac{-2}{\sqrt{30/5}}\right|$$

$$= \frac{2}{2.45}$$

$$= 0.816$$

**Expected value:**

$$t_e = \left|\frac{\bar{d}}{\sqrt{S^2/n}}\right|$$ follows t-distribution with $5-1=4$ d.f

$$= 4.604$$

**Inference:**

Since $t_0 < t_e$ at 1% level of significance we accept the null hypothesis. We therefore, conclude that there is no change in IQ after the training programme.