classmate

Date

Page

Module-6

Hypothesis Testing-IInd

small sample test: Student t- test, F- test- chi-square test - goodness of fit - independence of attributes - Design of experiments- Analysis of variance - One way- Two way- Three way Classifications- CRD - RBD - LSD.

If the sample size is large ($n \geq 30$), then the sampling distribution of statistic is normal. but if samples are small ($n \leq 30$) then the above result does not hold good.

For estimation of parameter as well as for testing a hypothesis, following distributions are used:-

(1) Student's t- distribution
(2) F- distribution
(3) Chi-square (χ^2) distribution

When the sample size is less than 30 ($n<30$), the sample distribution of many statistic is not normal, even though the parent population may be normal. Also the assumption of equality of population parameter and corresponding sample statistic will not be justified for small samples.

Degree of freedom:-

No of different parameter used used to calculated a statistic. Degree of freedom is often represented by r. Calculated as

[v = sample size − no. of restriction]

$$\boxed{r = n-k}$$

classmate
Date
Page

(1) Student's t- distribution :-

Let $x_1, x_2, \dots x_n$ be a random sample of size n $(n < 30)$, drawn from a normal population with mean μ and S.D σ. then student's t-statistic is defined by

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}}$$

$\bar{x}$ is sample mean and t is a random variable having t - distribution with $v = n-1$ degree of freedom and with p.d.f

$$f(t) = c\left(1 + \frac{t^2}{v}\right)^{-\frac{(v+1)}{2}}, \quad \text{where } v = n-1$$

c is a constant required to make area under the curve unity i.e

$$\boxed{\int_{-\infty}^{\infty} f(t)\,dt = 1}$$

t distribution is used
(1) when sample size is < 30.
(2) ~~Sample's~~ Population's S.D is not known

Properties of t- distribution :-

1. The t - distribution is asymtotic to the x axis ie its tends to infinity on either side.
2. The t - distribution is symmetrical about mean.

(3) Shape of the curve varies with degree of freedom.

(4) The larger the no. of degree of freedom, the more likely t - distribution resembles standard normal distribution.

(5) sampling distribution of t doesnot depend on population parameter but depends on degree of freedom (sample size).

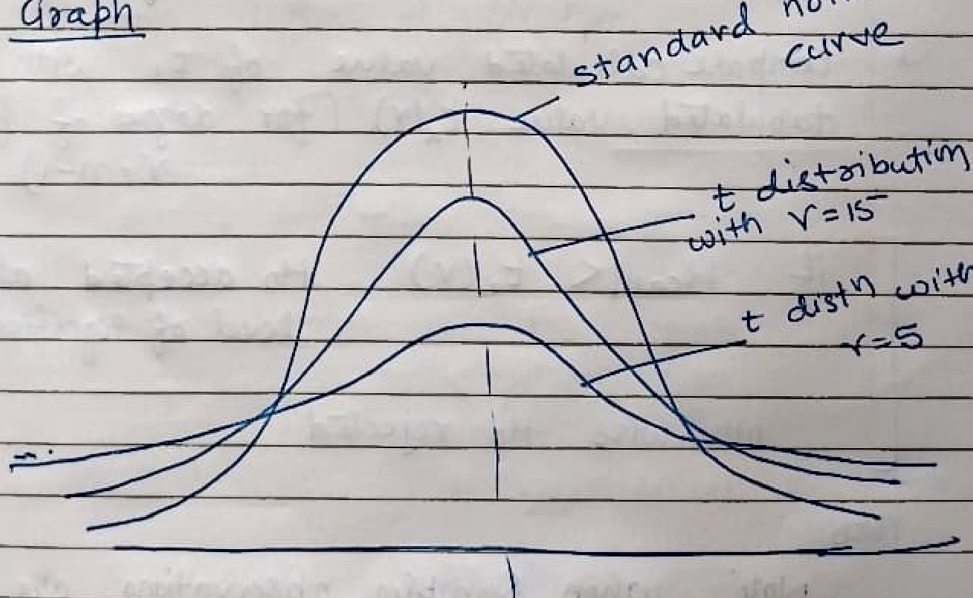
Uses of t-distribution:-

The t- distribution is used for.

(1) Test the significance of difference b/w. sample mean and population mean

(2) Test the significance of difference b/w. the mean of two small samples.

Graph:



(1) Test of significance for single mean:-

(ie test of significance of the difference b/w sample mean and population mean)

test statistic $\quad t = \dfrac{\bar{x} - \mu}{S/\sqrt{n-1}}$

$\bar{x}$ = sample mean
μ = mean of population
n = sample size
S = standard deviation.

* Degree of freedom $= n-1$ (as mean is used to calculate the test statistic ie $K=1$)
* if S is not given, it can be computed from the given data, using.

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} \quad \text{with} \quad \boxed{\nu = n-1}$$

Degree of freedom.

* Compare calculated value of t, with its tabulated value $t_\alpha(\nu)$ (for degree of freedom $\nu = n-1$)

[if $|t_{cal}| < t_\alpha(\nu)$, Ho accepted at α level of significance

otherwise Ho rejected]

Note:- when samples observations are given:-

$$\bar{x} = \frac{\Sigma x}{n} \qquad S = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2}$$

or $\quad \bar{x} = \dfrac{\Sigma d}{n} + A, \quad S = \sqrt{\dfrac{\Sigma d^2}{n} - \left(\dfrac{\Sigma d}{n}\right)^2}$, A is assumed mean.

Confidence limit :-

(i) 95% confidence limit for μ is given by

$$\left| \frac{\bar{x} - \mu}{s/\sqrt{n-1}} \right| \le t_{.05}$$

ie $$\boxed{\bar{x} - t_{.05} \frac{s}{\sqrt{n-1}} \le \mu \le \bar{x} + t_{.05} \frac{s}{\sqrt{n-1}}}$$

where $t_{.05}$ is the 5% critical value for $v = n-1$ degree of freedom for two tailed test.

(ii) 99% confidence limit

$$\bar{x} - t_{.01} \frac{s}{\sqrt{n-1}} \le \mu \le \bar{x} + t_{.01} \frac{s}{\sqrt{n-1}}$$

where $t_{.01}$ is the 1% critical value of t for $v = n-1$ degree of freedom.

Ex:- The mean monthly sale of a particular household item in departmental stores was observed as 13.3 unit per store. After an advertizing compaign, data on sales of this item from 25 stores were collected and it is found that the mean sales has increased to 141.5 unit with a standard deviation of 15.2 units. Was the advertising compaign successful? Assume that level of significance is 5%.

Solution:- $n = 25$, $\bar{x} = 141.5$ $S = 15.2$

$\mu = 133.3$ $\nu = n-1 = 24$

Null Hypothesis:- $H_0 : \mu = 133.3$

$H_1 : \mu > 133.3$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{141.5 - 133.3}{\frac{15.2}{\sqrt{25-1}}} = 2.665$$

Using table

$t_{.05}$ with $\nu = 24$ is 2.064

Now

as $|t_{cal}| > t_{.05}$

ie H_0 is rejected.

H_1 is accepted

ie Compaign was successful.

Ex 2: A machinist is making engine. engine parts with axle diameter of 0.7 inch. A random sample of 10 parts shows a mean diameter of 0.742 inch with a S.D of .04 inch. compute the statistic you would use to test whether the work is meeting the specification.

Sol $\mu = 0.7$ $\bar{x} = .742$ $S = .04$ $n = 10$ $v = n-1$
$v = 9$

Null Hypothesis

$H_0 : \mu = 0.7$. ie work is meeting the specification

Alternate

$H_1 : \mu \neq 0.7$

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}} = \frac{0.742 - 0.7}{\frac{.04}{\sqrt{10-1}}}$$

$t = 3.15$

Now

$t_{.05}$ with 9 degree of freedom = 2.262

as $|t_{cal}| > t_{.05}$

ie H_0 is rejected

ie H_1 accepted

ie work is not meeting the specification.

classmate

Date
Page

Ex3 A random sample of size 16 has 53 as mean. The sum of the squares of the deviation taken from mean is 135. Can this be regarded as taken from the population having 56 as mean. Obtain 95% and 99% confidence limit of the mean of the population.

sol^n:- $\bar{x} = 53$ $\quad n = 16$ $\quad \Sigma(x-\bar{x})^2 = 135$

$\mu = 56,$ $\quad S = \sqrt{\dfrac{\Sigma(x-\bar{x})^2}{16}} = \sqrt{\dfrac{135}{16}} = 2.904$

$H_0 : \mu = 56$

$H_1 : \mu \neq 56$ (two tailed)

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}} = -4$$

$t_{.05}$ for $\nu = n-1 = 15$ is 2.131

as $|t_{cal}| > t_{.05}$

ie H_0 is rejected.

ie sample has not come from a population having 56 as mean.

99% confidence limit for population mean

$$\bar{x} - \frac{S}{\sqrt{n-1}} \times t_{.05} \leq \mu \leq \bar{x} + \frac{S}{\sqrt{n-1}} t_{.05}$$

ie:- $\mu \in [51.402, 54.598]$

Slly 99% confidence limit

$\mu \in [\bar{x} \pm \frac{s}{\sqrt{n}} t_{.01}]$ $\boxed{t_{.01} = 2.947}$

$\mu \in [50.71, 55.21]$

Q The mean life of sample of 25 bulbs is found as 1550 hours with a S.D of 120 hours. The company manufacturing the bulbs claims that the average life of their bulbs is 1600 hours. Is the claim accepted as 5% level of significance.

Soln:- $n = 25$, $\bar{x} = 1550$ $S = 120$ $\mu = 1600$

Null Hypo $H_0 : \mu = 1600$ hours, ie average life of bulbs is 1600 hours

Alternate Hypo:
$H_1 : \mu < 1600$ (one tailed)

$\alpha = .05$ (5%)
degree of freedom $v = 25-1 = 24$

$t_\alpha = 1.711$

test statistic

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} = \frac{1550 - 1600}{120/\sqrt{24}} = \frac{-50\sqrt{24}}{120}$$

$$= \frac{-50 \times 4.898}{120}$$

$$= -2.0412$$

as $|t_{cal}| > t_\alpha$
ie Null Hypo is rejected. H_1 is accepted

classmate
Date
Page

Ex:- The life time of electric bulbs for a sample of 10 from a large consignment gave the following data:

item	1	2	3	4	5	6	7	8	9	10
life in ('000) hours	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6

can we accept the hypothesis that the average life time of bulb is 4000 hours.

solⁿ:- Let $H_0 : \mu = 4000$ ie there is no significant difference b/w sample and population mean.

first calculate sample mean and S.D

x	$x-\bar{x}$	$(x-\bar{x})^2$
4.2	−0.2	.04
4.6	0.2	.04
3.9	−0.5	0.25
4.1	−0.3	0.09
5.2	0.8	0.64
3.8	−0.6	0.36
3.9	−0.5	0.25
4.3	−0.1	0.01
4.4	0	0
5.6	1.2	1.44
$\Sigma x = 44$		$\Sigma(x-\bar{x})^2 = 3.12$

$\bar{x} = \frac{\Sigma x}{n} = 4.4$

$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} =$

Let $H_0 : \mu = 4$ (in 000 hours)

$H_1 : \mu \neq 4$ (")

(two tailed)

$S = \sqrt{\frac{3.12}{10}} = 0.5585$

test statistic

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} = \frac{4.4 - 4}{.5585/\sqrt{9}}$$

$$t = \frac{.4 \times 3}{.5585} = 2.148$$

for $\nu = n-1 = 9$

$t_\alpha = 2.262$ (Let $\alpha = 5\%$)

as $|t_{cal}| < t_\alpha$ ie Null Hypo is accepted.
ie we accept the hypo. that the average life time of bulbs is 4000 hours.

Q Test made on the breaking strength of 10 pieces of a metal wire gave the results 578, 572, 570, 568, 572, 570, 570, 572, 596, 584. Test if the mean breaking strength of the wire can be assumed as 577 kg.

Soln

x_i	$d_i = x_i - A$	d_i^2
578	8	64
572	2	4
570	0	0
568	-2	4
572	2	4
570	0	0
570	0	0
572	2	4
596	26	676
584	14	196
	52	952

Let A = 570

$n = 10$.

$$S = \sqrt{\frac{\Sigma d_i^2}{n} - \left(\frac{\Sigma d_i}{n}\right)^2}$$

$$S = \sqrt{\frac{952}{10} - \left(\frac{52}{10}\right)^2}$$

$$S = \sqrt{95.2 - (5.2)^2}$$

$$\boxed{S = 8.255}$$

$$\bar{x} = A + \frac{\Sigma d}{n} = 570 + \frac{52}{10}$$

$$= 570 + 5.2 = 575.2$$

$H_0: \mu = 577$

$H_1: \mu \neq 577$ (two tailed test)

Now

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{572.2 - 577}{\frac{8.255}{\sqrt{9}}}$$

$$= \frac{-4.8 \times 3}{8.255}$$

$$= -1.744$$

Now t_α at 5% level of significance for $\nu = 9$ is

$$t_\alpha = 2.262$$

as $|t_{cal}| > t_\alpha$

ie H_0 is rejected

ie Mean breaking strength of wire can not be assumed as 577.

classmate

Date
Page

Test 2 :- Test for difference of Means :-

If two samples of size n_1 and n_2 with means $\bar{x}$ and $\bar{y}$ and S.D s_1 and s_2 are drawn from a normal population, having same S.D.

The student's t - statistic is given by

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{——} \quad (1)$$

where $S = \sqrt{\dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$ $\qquad V = n_1 + n_2 - 2$

case I :- When $n_1 = n_2 = n$ and samples are independent ie observation in two samples are not correlated then test statistic is given by :

from (1)

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{ns_1^2 + ns_2^2}{2n-2}}\sqrt{\frac{1}{n} + \frac{1}{n}}}$$

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{n}{2}} \times \sqrt{\frac{2}{n}}\sqrt{\frac{s_1^2 + s_2^2}{n-1}}}$$

ie $$\boxed{t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2 + s_2^2}{n-1}}}}$$ with $V = 2n - 2$

Case II :- If $n_1 = n_2 = n$ and if the pairs of values of x and y are associated or correlated in some way (not independent) then above formula can'nt be used.

Let $d_i = x_i - y_i$ denote the difference (with sign) in the values of x_i & y_i

then test statistic is

$$\boxed{t = \frac{\bar{d}}{\frac{S}{\sqrt{n-1}}}} \quad \text{with } \boxed{\nu = n-1}$$

where $\bar{d} = \frac{\Sigma d}{n}$ and $S = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$

Q. The means of two random samples of size 9 and 7 are 196.42 and 198.82 respectively. The sum of squares of the deviation from the mean are 26.94 and 18.73 resp. Can the sample be considered to have been drawn from same population?

Solⁿ. $n_1 = 9$, $n_2 = 7$

$\bar{x}_1 = 196.42$, $\bar{x}_2 = 198.82$

$\Sigma(x_1 - \bar{x}_1)^2 = 26.94$, $\Sigma(x_2 - \bar{x}_2)^2 = 18.73$

$S_1 = \sqrt{\dfrac{\Sigma(x_1-\bar{x}_1)^2}{n_1}}$ $\qquad S_2 = \sqrt{\dfrac{\Sigma(x_2-\bar{x}_2)^2}{n_2}}$

$S_1 = \sqrt{\dfrac{26.94}{9}}$ $\qquad S_2 = \sqrt{\dfrac{198.82}{n_2}}$

$$S = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1+n_2-2}} = \sqrt{\frac{26.94 + 18.73}{9+7-2}}$$

$$\boxed{S = 1.806}$$

Null Hypothesis $H_0 : \mu_1 = \mu_2$ ie Sample are drawn from the same population

Alternate Hypothesis :- $H_1 : \mu_1 \neq \mu_2$ (Two tailed Test)

$\alpha = .05$ (assume)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{196.42 - 198.82}{1.8061\sqrt{\frac{1}{9} + \frac{1}{7}}}$$

$$= -2.6368$$

$$|t| = 2.6368$$

$$v = n_1 + n_2 - 2$$

$$\boxed{v = 9 + 7 - 2 = 14}$$

from the table $t_\alpha (v = 14) = 2.145$

as $|t| > |t_\alpha|$

ie Null Hypo. is rejected. ie sample are not drawn from same population.

Q 2 The mean height and SD height of 8 randomly chosen soldiers are 166.9 cm and 8.29 cm respectively. The corresponding values of 6 randomly chosen sailors are 170.3 cm and 8.50 cm resp. Based on this data, can we conclude that soldiers are in general shorter than

classmate

Date ____
Page ____

Soln:- $n_1 = 8,\ n_2 = 6,\ \overline{x}_1 = 166.9$ cm $\quad \overline{x}_2 = 170.3$ cm

$s_1 = 8.29$ cm $\quad s_2 = 8.50$ cm

$$S = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

$$\boxed{S = 9.05 \text{ cm}}$$

Null Hypothesis $H_0 : \mu_1 = \mu_2$ i.e. there is no significance difference b/w the heights of soldiers and sailors.

Alternate Hypothesis :- $H_1 : \mu_1 < \mu_2$ (one tailed)

$\alpha = .05$ (assume)

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

$$t = \frac{166.9 - 170.3}{9.05\sqrt{\frac{1}{8} + \frac{1}{6}}}$$

$t = -.696$

$|t| = .696$

Critical value :- $\quad v = n_1 + n_2 - 2 = 12$

ie $t_\alpha (v = 12) = 1.782$

as $|t| < t_\alpha$ ie H_0 is accepted.

ie there is no significant difference b/w height of soldiers and sailors and we cannot conclude that the sailors are in general shorter than sailors.

Ex3 Samples of two types of electric bulbs were tested for length of life and the following data were obtained

	Size	Mean	S.D
Sample 1	8	1234 hrs	36 hrs
Sample 2	7	1036 hrs	40 hrs

Is the difference in the means is sufficient to warrant that type I bulbs are superior to type 2.

Solⁿ:- $n_1 = 8$, $n_2 = 7$, $\overline{x_1} = 1234$ hrs

$\overline{x_2} = 1036$ hrs.

$s_1 = 36$ hrs $\quad s_2 = 40$ hrs

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}} = 40.73 \text{ hrs}$$

Null Hyp.: $H_0: \mu_1 = \mu_2$

Alternate Hyp.: $H_1: \mu_1 > \mu_2$ (one tailed)

$\alpha = .05$ (assume)

test statistic:-

$$t = \frac{\overline{x_1} - \overline{x_2}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{1234 - 1036}{40.73\sqrt{\frac{1}{8} + \frac{1}{7}}}$$

$$t = 9.39$$

$|t| = 9.39$

Critical value: degree of freedom $\nu = n_1 + n_2 - 2$
$\nu = 13$

$t_{.05}(\nu = 13) = 1.771$

as $|t| > t_{.05}(\nu = 13)$

ie H_0 rejected ie type I bulb are superior to type 2 bulbs.

Q:- A group of 5 patients treated with medicine A, weigh 42, 39, 48, 60, 41 kg. Second group of 7 patients from the same Hospital treated with medicine B, weigh 38, 42, 56, 64, 68, 69 and 62 kg. Do you agree with the claim that medicine B ~~is~~ increases the weigh significantly.

(Sol^n:- medicine A and B donot differs significantly as regards their effect on increase in weight)

Sample size are equal:- $(n_1 = n_2)$

Q:- The following data represent the biological values of protein from cow's milk and buffalo's milk at a certain level.

cow's milk :	1.82	2.02	1.88	1.61	1.81	1.54
Buffalo's milk	2.00	1.83	1.86	2.03	2.19	1.88

examine if the average values of protein in the two samples significantly differ.

classmate
Date
Page

Solⁿ: $n_1 = n_2 = 6$ (Sample size is equal)
& two samples are independent

$n=6$, $\bar{x_1} = 1.78$, $\bar{x_2} = 1.965$
$S_1 = .16$, $S_2 = .124$ (calculate from given data)

Null Hypo:- H_0: $\mu_1 = \mu_2$ ie there is no significant difference in average values of proteins in two milk samples.

Alternate Hypo:- H_1: $\mu_1 \neq \mu_2$ (two tailed)

$\alpha = .05$ (assume)

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2 + s_2^2}{n-1}}} = \frac{1.78 - 1.965}{\sqrt{\frac{(0.16)^2 + (0.124)^2}{6-1}}}$$

$$t = -2.043$$

$$|t| = 2.043$$

critical value:-

$$v = n + n - 2 = 2n-2$$
$$\boxed{v = 10}$$

$$t_{.05}(v=10) = 2.228$$

as $|t| < t_{.05}$ ie null Hypo. is ~~reject~~ accepted.

classmate

Date
Page

ie there is no. significant difference b/w the average values of proteins in two milk samples.

Q:- A certain injection administered to 12 patient resulted in the following changes of blood pressure:-

5, 2, 8, -1, 3, 0, 6, -2, 1, 5, 0, 4

can it be concluded that the injection will be in general accompanied by an increase in blood pressure.

solⁿ:- Here the $d = \bar{x}_1 - \bar{x}_2$ is given ie change in the blood pressure.
Let x_2 is initial blood pressure and x_1 is final blood pressure.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$

[Samples are correlated (not independent)]

$n = 12, \ \Sigma d = 31 \qquad \Sigma d_i^2 = 185$

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{31}{12} = 2.58$$

$$S = \sqrt{\frac{\Sigma d_i^2}{n} - \left(\frac{\Sigma d_i}{n}\right)^2} = \sqrt{\frac{185}{12} - \left(\frac{31}{12}\right)^2}$$

$$\boxed{S = 2.96}$$

$\alpha = .05$

$$t = \frac{\bar{d}}{\frac{S}{\sqrt{n-1}}} = \frac{2.58}{\frac{2.96}{\sqrt{12-1}}} = \cancel{2.79}\ 2.89$$

~~Nullt~~

Critical value: $\nu = n-1 = 11$ (degree of freedom)

$$t_{.05}(\nu=11) = 1.796$$

as $|t| > t_{.05}$

ie null Hypo. is rejected - ie injection will in general accompanied by an increase by an increase in blood pressure.

Q:- Scores obtained in a shooting competition by 10 soldiers before and after training are given below:-

Score before training:	67	24	57	55	63	54	56	68	33	43
" after " :	70	38	58	58	56	67	68	75	42	38

Test whether the intensive training is useful at .05 level of significance.

Soln.

$n_1 = n_2 = 10$ are samples are correlated. (not independent)

$d = x-y$:	-3	-14	-1	-3	7	-13	-12	-7	-9	5
d^2 :	9	196	1	9	49	189	144	49	81	25

$$\bar{d} = \frac{\Sigma d}{n} = \frac{-50}{10} = -5$$

$$s = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2} = 6.44$$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 < \mu_2$ (one tailed)

test statistic :-

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n-1}}} = \frac{-5}{\frac{6.94}{\sqrt{9}}} = -2.16$$

Critical value :-

$t_\alpha(9) = 1.96$

degree of freedom $\nu = n-1$
$= 9$

as $|t| > t_\alpha$

is H_0 is rejected. ie intensive training is useful.