Unit-3 Testing of Hypothesis.

Symbols :-

	Population	Sample
Size	N	n
Mean	μ	$\bar{x}$
S.D	σ	s
Proportion	P	p

Problems under large samples : ($n > 30$)
we use (z test)

1. A sample of 900 members has mean 3.4 and standard deviation 2.61. Is a sample from a large population of mean 3.25 and S.D = 2.61. Assuming population is normal. Find 95% confidence limits for its mean.

95% acceptable

5% reject

$n = 900$

$\bar{x} = 3.4$

$s = 2.61$

$\mu = 3.25$ $\qquad \sigma = 2.61$

Using

$$\boxed{z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}}$$

For all problems we use $H_0 \to$ null hypo

$H_1 \to$ alternate hypo

Null Hypothesis H_0:

$\bar{x} = \mu$

Alternate Hypothesis is H_1:

$\bar{x} \neq \mu$

$$z \text{ test} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{3.4 - 3.25}{2.61/\sqrt{900}}$$

$$= 1.724$$

Calculated value : 1.724

Table Value = 1.96 (5% level of significance)

calculated value < table value

$\therefore$ Accept H_0.

P:2 The main yeild of wheat from a district A was 210 pounds with SD. 10 pounds per acre from a sample of 100 plots In another district B main yeild was 220 po S.D 12 pou from sample of

150 plots. Assuming that S.D of entire State was 11 pounds. Main yield of crops in 2 samples are 5% of LOS.

$\bar{x}_1 = 210, \; S_1 = 10, \; n_1 = 100$

$\bar{x}_2 = 220, \; S_2 = 12, \; n_2 = 150$

$\sigma = 11$

$\sigma = 11$

$H_0 : \bar{x}_1 = \bar{x}_2$

$H_1 : \bar{x}_1 \neq \bar{x}_2$

$$|z| = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} = \frac{210 - 220}{\sqrt{\frac{121}{100} + \frac{121}{150}}}$$

$=$

Table value at 5%

$= 1.96$

The means of 2 large samples 1000 and 2000 members are 67.5 inches and 68 inches resp. Can the Samples be regarded as drawn from the same population of S.D 2.5 inches.

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$$

$H_0 = \bar{x}_1 = \bar{x}_2$

$H_1 = \bar{x}_1 \neq \bar{x}_2$

$n_1 = 1000$, $n_2 = 2000$

$\bar{x}_1 = 67.5$, $\bar{x}_2 = 68$

$\sigma = 2.5$

$|z| = |-5.1546|$

$= 5.1546$

calculated z-value = 5.154

Table value at 5% level of significance = 1.96

If the significance is not given in question, take it as 5% as default

In a college, 60 Juniors are found to have a mean height of 171.5 cm, 50 senior students are found to have mean height of 173.8 cm. Can it be concluded that the juniors are shorter than seniors at 5% level of significance. assuming S.D of clg is 6.2cm

Sol:

$n_1 = 60$; $n_2 = 50$

$\bar{x}_1 = 171.5$ $\quad \bar{x}_2 = 173.8$

$\sigma = 6.2$

H_0: $\bar{x}_1 < \bar{x}_2$ (Juniors are lesser in height than seniors) [Right tailed test]

H_1: $\bar{x}_1 = \bar{x}_2$

Z-calculated = 1.937

Table value = 1.645

H_0 is rejected because calculated value is greater than table value.

Average marks scored by 32 boys is 72 with S.D of 8. for 36 girls is 70 with S.D of 6 and 1% level of significance. Whether boys perform better than girls?

\# $n_1 = 32, \; n_2 = 36$

$\bar{x}_1 = 72, \; \bar{x}_2 = 70$

$s_1 = 8, \; s_2 = 6$

$H_0 = \bar{x}_1 > \bar{x}_2$

$H_1 = \bar{x}_1 \ngtr \bar{x}_2$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$|z| = 1.154$

Table value at 1%.

Unit-3 Definitions

Sampling: A part selected from a population is sample & its selection process is called sampling.

Random sampling: One in which each member of population has equal chance of being included in it.

Standard error: is S.D of sampling distribution

Test of significance:

1) The deviation between observed sample statistic and hypothetical parameter value

2) The deviation between two sample statistic

Unit-3 continuation

Test for the significant difference between sample proportion & population proportion

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} \qquad Q = 1 - P$$

Test for the significant diff between 2 sample proportion

(i) P is known

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

(ii) P is unknown

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

P: A sample of $\overset{n_1}{\underline{400}}$ proportion of Tea drinkers is $\overset{p_1}{\underline{0.0125}}$ and in another sample of $\overset{n_2}{\underline{1200}}$, the proportion is $\overset{p_2}{\underline{0.0083}}$. Test whether the samples are taken from population which has

Proportion (cap. P) → population prop. 0.01.

Sol:

No significant difference H_0 = betn 2 sample prop

$\boxed{Q = 1 - 0.01}$

$$\begin{array}{r} 1.00 \\ -0.01 \\ \hline 0.99 \end{array}$$

signifi. diff betn 2 sam H_1.

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.0125 - 0.0083}{\sqrt{0.01 \times 0.99\left(\frac{1}{400} + \frac{1}{1200}\right)}} \quad = 0.0042$$

$$= 0.73$$

$0.0099 \times 0.003 = 0.0000297$

= Take level of significance as 5%.

1%	5%	10%
2.58	1.96	1.645

Table value of 5% is 1.96

$0.73 < 1.96$

H_0 is accepted

A Random sample of 400 men and 600 women were asked if they would like to have a flyover near their residence. 200 men & 325 women were in favor of it. Test the equality of proportion of men & women in this proposal.

~~Sol~~

$n_1 = 400, \quad n_2 = 600$

$p_1 = \frac{200}{400}, \quad p_2 = \frac{325}{600}$

Ho : no significant diff btn men & women proposal.

H_1 : significant diff btn.

P is unknown

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}.$$

$P = 0.524$

$Q = 1 - 0.524 = 0.476$

$$Z = \frac{\frac{200}{400} - \frac{325}{600}}{\sqrt{(0.524 \times \cancel{2})(0.476)\left[\frac{1}{400} + \frac{1}{600}\right]}}$$

$$= \frac{\frac{1}{2} - \frac{325}{600}}{\sqrt{(0.524)(0.476)\left[\frac{1}{400} + \frac{1}{600}\right]}}$$

$|z| = 1.25$

Table value at 5% los is 1.96

c.v < table value H_0 accepted.

If $n \geq 30$ z test

$n < 30$, t test.

Small Sample test

i) t test

Type 1: To find is there any significant difference between sample mean & population mean

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \quad \text{where } s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}, \ \bar{x} = \frac{\Sigma x}{n}$$

Degrees of freedom $= n - 1$
(D.o.f)

Type 2:-

To find there is any significant diff between 2 sample means.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad \text{where } s = \sqrt{\frac{\Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}}$$

D.O.F $= n_1 + n_2 - 2$

Type 3

~~t-test~~ for Paired observation (here we Consider 2 independent samples Such that there are situations where the pair of values of x_1 & x_2 are corelated).

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

$$\bar{d} = \frac{\Sigma d}{n} \qquad d = x_1 - x_2$$

$$D.O.F = n-1$$

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

P:1 The mean weekly sales of a certain Talcum powder in a large group of departmental store was 146.3 tins per store. After an ad campaign, the mean weekly sales of 22 stores for a typical week increased to 153.7 and showed a S.D of 17.2. Was the ad campaign successful?

$n = 22$ population mean $= 146.3$

Sample mean $= 153.7$

Sample SD $= 17.2$

H_0: There is no significant difference

H_1: There is a significant difference

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

$t = 1.97$

Table value ~~of~~ at 5% level of significance with $(n-1)$ d.o.f (i.e) $(22-1 = 21$ d.o.f$)$ is 2.08

cal value < table value

$\therefore H_0$ is accepted.

:2 A random sample of 10 boys have the following IQs

x
70
120
110
101
88
83
95
98
107
100

Do this data support the assumptions of a population mean IQ is 100.

Sol:

x	$(x-\bar{x})$	$(x-\bar{x})^2$	$\Sigma(x-\bar{x})^2$
70	$(70-97.2)=-27.2$	739.84	
120	$(120-97.2)=22.8$	519.84	
110	12.8	163.84	
101	3.8	14.44	
88	-9.2	84.64	1833.6
83	-14.2	201.64	
95	-2.2	4.84	
98	0.8	0.64	
107	9.8	96.04	
100	2.8	7.84	

$$\bar{x} = \frac{\Sigma x}{n} = \frac{972}{10} = 97.2$$

$$\mu = 100, \quad n = 10$$

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

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

$$= \sqrt{\frac{1833.6}{10}}$$

$$= 183.36$$

$$t = \frac{\bar{x}-\mu}{S/\sqrt{n-1}} = \frac{97.2-100}{183.36/\sqrt{10-1}} = 0.62$$

Table value at 5% los with (10-1=9) dof is 2·26

Calculated value < Table value

H_0 is accepted.

Two Horses A and B were tested according to the time to run a particular race with following result

Horse A x_1	Horse B x_2
28	29
30	30
32	30
33	24
33	27
29	29
34	—

$\bar{x}_1 > \bar{x}_2$

1 tail

Test whether horse A is running faster than horse B at 5% los.

sol:

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

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

$n_1 = 7 \quad n_2 = 6$

x_1	x_2	$(x_1-\bar{x}_1)$	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)$	$(x_2-\bar{x}_2)^2$
28	29	-3.28	0.84	.10.75	0.705
30	30	-1.28	1.84	1.63	3.38
32	30	0.72	1.84	0.518	3.38
33	24	1.715	-4.16	2.94	17.3
33	27	1.715	-1.16	2.94	1.345
29	29	-2.25	0.84	5.06	0.705
34	-	2.715		7.37	

$$\bar{x}_1 = \frac{\Sigma x}{n} = \frac{219}{7} = 31.285$$

$$\bar{x}_2 = \frac{\Sigma x_2}{n} = \frac{169}{6} = 28.166$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = 2.71$$

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

$\boxed{s = 2.71}$

Table value at 5% los with (n_1+n_2-2) of dof is 1.796.

The nicotin contents in mg of two samples of tobacco were found to be as follows

A	B
x_1	x_2
24	27
27	30
26	28
21	31
25	22
-	36

can it be said that 2 samples came from same normal population

H_0: There is no significant btn 2 samples

H_1: There is "

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

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

$\overline{x}_{12} =$

A	B	$(x_1-\overline{x}_1)^2$	$(x_2-\overline{x}_2)^2$
x_1	x_2		
24	27	-0.36	4
27	30	5.76	1
26	28	1.96	+1
21	31	12.96	4
25	22	0.16	49
-	36		49
		21.2	108

$\overline{x}_1 = \frac{123}{5}$
$= 24.6$

$n_1 = 5 \quad n_2 = 6$

$\overline{x}_2 = \frac{174}{5}$
$= 29$

$$S^2 = \frac{21.2 + 108}{5+6-2}$$

$S = 3.78$

$$t = \frac{24.6 - 29}{3.78\sqrt{\frac{1}{5}+\frac{1}{6}}} = \frac{24.6-29}{3.78\sqrt{\frac{11}{30}}}$$

$$= \frac{-4.4}{3.78 \times 0.1}$$

$|t| = 3.4$

Table value at 5% los for n_1+n_2-2 d.o.f

9 dof 2.26

H_0 is accepted.

C.V < T.V

t test for paired observation.

~~t test~~.

IQ test were determined for 5 people before and after they were trained the result are as follows

Before	After
110	120
120	118
123	125
132	136
125	121

Do the data support the training.

H_0: There is no significant difference.

H_1: There is significant difference

$$t = \frac{\bar{d}}{s/\sqrt{n-1}} \quad ; \quad \bar{d} = \frac{\Sigma d}{n}$$

$$s^2 = \frac{\Sigma d^2}{n} - (\bar{d})^2$$

Before	After	$d = x_1 - x_2$	d^2
110	120	-10	100
120	118	2	4
123	125	-2	4
132	136	-4	16
125	121	4	16
		$\Sigma d = -10$	140

$$t = \frac{\bar{d}}{s/\sqrt{n-1}} \; ; \quad \bar{d} = \frac{\Sigma d}{n} = \frac{-10}{5} = -2$$

$$s^2 = \frac{\Sigma d^2}{n} - (\bar{d})^2 = 24$$

$$\boxed{s = 4.89}$$

$$t = \frac{-2}{4.89/\sqrt{5-1}}$$

$$t = 0.817$$

Table value at 5% los with $(n-1 = 4)$ d.o.f = ~~2.77~~ 2.776.

2.4