

## "IV. TESTING OF HYPOTHESIS."

\* **Population** :- It is a collection of objects.

- The no. of objects in the population is called size of the population.
- Population may be finite or infinite.

Ex:- The no. of students in a college is a finite population.

The products produced in a factory can be considered as an infinite population.

\* **Sample** :- It is a subset of population. There

are 2 types of samples.

1) Small Sample.

2) Large Sample.

1. **Small Sample** :- When the size of the sample ( $n$ ) ( $n < 30$ ) then the sample is called as a sample size.

2. **Large Sample** :- When the size of the sample ( $n \geq 100$ ) then it is called as large sample.

\* **Testing of hypothesis** :-

It is a process for deciding whether to accept or to reject the hypothesis is called testing of hypothesis.

There are 2-types of hypothesis.

1. Null hypothesis ( $H_0$ )

2. Alternative hypothesis ( $H_1$ ).

1. Null hypothesis :- To decide whether one procedure is better than another procedure we from (is better than) the hypothesis ie no significant difference blw. the procedures such hypothesis is called as Null hypothesis. It is denoted as  $H_0$ .

2. Alternative hypothesis :- Any hypothesis which is complement to the Null hypothesis is called alternative hypothesis.

- It is denoted as  $H_1$ .

\* Two-tailed test :- In a testing of hypothesis whether the alternative hypothesis two-tailed test, we assume that  $\rightarrow$  Null hypothesis  $H_0: \mu = \mu_0$ .

Against the alternative hypothesis.

$$H_1: \mu \neq \mu_0$$

\* Right - tailed test :-

$$\text{Null hypothesis } H_0: \mu = \mu_0$$

Against the alternative hypothesis  $H_1: \mu > \mu_0$

\* Left - tailed test :-

$$\text{Null hypothesis } H_0: \mu = \mu_0$$

Against the alternative hypothesis  $H_1: \mu < \mu_0$

## Errors of Sampling :-

There are 2-types of errors.

① Type 1 error (or)  $\alpha$  error.

② Type 2 error (or)  $\beta$  error.

① Type 1 error:- Reject  $H_0$  when it is true. If the Null hypothesis  $H_0$  is true but it is rejected by test procedure then, the error is called as type 1-error.

② Type 2-error:- Accept  $H_0$  when it is false. If the Null hypothesis is false, but it is accepted by test procedure then the error is called as type 2-error.

$Z_{\text{tab}} > Z_{\text{crit}}$  accept.

$Z_{\text{tab}} < Z_{\text{crit}}$  reject.

\* Level of Significance & Confidence level.

Table values:-

Tail of test	level of significance ( $\alpha$ ) & $Z_\alpha$ .			
	1%	2%	5%	10%
Two-tailed test	2.56	2.32	1.96	1.64
Right - tailed test	2.32	1.96	1.64	1.26
left - tailed test	-2.32	-1.96	-1.64	-1.26

- If  $z$  table value is greater than  $z$  calculated value we accept Null hypothesis ( $H_0$ ).  
 → If  $z$  table value is less than  $z$  calculated value we reject Null hypothesis ( $H_0$ )

① When a sample of 60 workers the average time is taken by them to get to work as 33.8 min with standard deviation 6.1 min can we reject the Null hypothesis  $H_0: \mu = 32.6$  in favour of alternative hypothesis  $H_1: \mu > 32.6$  at level of significance 1%.

$$\bar{X} = \text{Sample mean} = 33.8,$$

$$\mu = \text{Population mean} = 32.6$$

$$\sigma = \text{S.D} = 6.1$$

$$n = 60.$$

- ① Null hypothesis ( $H_0$ ) =  $\mu = 32.6$  } Right  
 ② Alternative hypothesis ( $H_1$ ) =  $\mu > 32.6$  } Tailed test.  
 ③ Level of significance :  $\alpha = 1\%$

$$z_{\text{cat}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{33.8 - 32.6}{\frac{6.1}{\sqrt{60}}}$$

$$\text{and } z_{\text{tab}} \text{ value for } 1\% = 1.523.$$

$$z_{\text{tab}} = 2.32.$$

$z_{\text{tab}} > z_{\text{cat}}$  we accept Null hypothesis.

② A sample of 400 products is taken from a population with standard deviation is "10". The mean of sample is 40. Test whether the sample has come from a population with mean?

$$\bar{x} = \text{sample mean} = 40$$

$$\mu = \text{population mean} = 38$$

$$\sigma = \text{standard deviation} = 10$$

$$n = 400$$

$$1) \text{ Null hypothesis } (H_0) : \mu = 38$$

$$2) \text{ Alternative hypothesis } (H_1) : \mu \neq 38$$

$$3) \text{ level of significance } (\alpha) : \alpha = 5\%$$

$$|z_{\text{cat}}| = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \Rightarrow \frac{40 - 38}{\frac{10}{\sqrt{400}}} = 4$$

$$z_{\text{tab}} = 1.96$$

$$\therefore z_{\text{cat}} > z_{\text{tab}} \text{ we reject } (H_0).$$

③ According to the NORM's established for a mechanical aptitude test person who are 80 year old have average height is 73.5 with S.D. 8.6. If 40 randomly selected persons of their average is 76.7 test the hypothesis at  $\mu = 73.2$  against the alternative hypothesis  $\mu = 73.2$  at the 1% of level of significance.

$$\mu = \text{population mean} = 73.2$$

$$\sigma = 8.6, n = 40, \mu = 73.2 \text{ and } \bar{x} = 76.7$$

$$(i) H_0 : \mu = 73.2 \quad (\text{Right - tailed test})$$

$$(ii) H_1 : \mu > 73.2$$

$$(iii) \alpha = 1\%$$

$$|Z_{\text{cat}}| = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{76.7 - 73.2}{\frac{8.6}{\sqrt{40}}} = 2.57$$

$|Z_{\text{tab}}| = 2.32$

$$|Z_{\text{tab}}| = 2.32$$

$\therefore Z_{\text{tab}} < Z_{\text{cat}}$  we reject the  $(H_0)$ .

- ④ A sample of nine hundred members has mean of 3.4 cm and S.D 2.61 cm is the sample has been taken from a large population of mean 3.25 cm. If the population is normal & '1' mean is unknown and also calculated 95% of confidence interval for the population.

$$\sigma = 2.61 \text{ cm}, n = 900, \mu = 3.25$$

$$\bar{x} = 3.4$$

$$1) H_0 : \mu = 3.25 \quad (\text{two-tailed test})$$

$$2) H_1 : \mu \neq 3.25$$

$$3. \alpha = 5\%$$

$$|Z_{\text{tab}}| = 1.96$$

$$|Z_{\text{cat}}| = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{3.4 - 3.25}{\frac{2.61}{\sqrt{900}}} = 1.72$$

$\therefore Z_{\text{tab}} > Z_{\text{cat}}$  we accept the  $(H_0)$ .

\* Confidence interval :-

$$\left[ \bar{x} + z_{\alpha} \left( \frac{\sigma}{\sqrt{n}} \right), \bar{x} - z_{\alpha} \left( \frac{\sigma}{\sqrt{n}} \right) \right]$$

$$x = 3.4, z_{\alpha} = \alpha = 1.96 \quad (\text{from table})$$

$$\left[ 3.4 + 1.96 \left( \frac{2.61}{\sqrt{900}} \right), 3.4 - 1.96 \left( \frac{2.61}{\sqrt{900}} \right) \right]$$

$$\Rightarrow (3.57, 3.22)$$

⑤ An ambulance service claims that it takes on the average less than 10 min to reach its destination in emergency calls. A sample of 36 calls has a mean of 11 min and variance of 16 mins. Test the claim at 0.05 level of significance.

$$\alpha = 5\%, \sigma^2 = 16, \sigma = \sqrt{16} = 4.$$

$$n = 36, \mu = 10, \bar{x} = 11.$$

i.  $H_0 : \mu = 10$  (left-tailed test).

$H_1 : \mu < 10$

$$\alpha_T = 5\%$$

$$|z_{\text{tab}}| = 11 - 10 = 1$$

$$= 1.644$$

$$|z_{\text{cat}}| = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{11 - 10}{\frac{4}{\sqrt{36}}} = 1.5$$

$\therefore z_{\text{tab}} > z_{\text{cat}}$  we accept ( $H_0$ ).

Test of significance difference of mean of 2

large samples :-

Let  $\bar{x}_1$  be the mean of sample of size  $n_1$  with population mean  $\mu_1$  and variance  $\sigma^2_1$ .

Let  $\bar{x}_2$  be the mean of sample size of  $n_2$  with population mean  $\mu_2$  and variance  $\sigma^2_2$ .

To test the significance difference of mean is given by.

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

If samples have been drawn from the same population.

$$\sigma^2_1 + \sigma^2_2 = \sigma^2$$

The test formula is

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

- ① If survey of buying habits 400 women shopers are chosen at random in super market 'A' located in a certain section of the city their average weekly food expenditure is Rs 250 with s.d of Rs 400 for 400 women shopers chosen at random in super market B in another section of the city the average weekly food expenditure is Rs 220 with s.d of Rs 55 test at 1% level of significance whether the

whether the average weekly food expenditure of population of shopers are equal.

$$n_1 = 400, \bar{x}_1 = 250, \sigma_1 = 40$$

$$n_2 = 400, \bar{x}_2 = 220, \sigma_2 = 55$$

i)  $H_0: \mu_1 = \mu_2$  (two-tailed test)

ii)  $H_1: \mu_1 \neq \mu_2$

iii)  $\alpha = 1\%$

$$Z_{tab} = 2.56$$

$$Z_{cat} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{250 - 220}{\sqrt{\frac{(40)^2}{400} + \frac{(55)^2}{400}}} = 8.82$$

$\therefore Z_{tab} > Z_{cat}$  we reject the  $(H_0)$ .

② The mean of 2 large samples of size 's' 1000, 2000 members are 67.5 inches & 68.5 inches respectively can the samples be regarded as drawn from the same population of s.d 2.5 inches.

$$n_1 = 1000, \bar{x}_1 = 67.8, \sigma = 2.5$$

$$n_2 = 2000, \bar{x}_2 = 68$$

i)  $H_0: \mu_1 = \mu_2$  (two-tailed test)

ii) Alternative  $H_1: \mu_1 \neq \mu_2$

iii)  $\alpha: \alpha = 5\%$

$$|Z_{tab}| = 1.96$$

$$|Z_{tab}| = |-5.16| = 5.16 \quad (\times)$$

$$|Z_{cat}| = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} = \frac{67.5 - 68}{\sqrt{\frac{(2.5)^2}{1000} + \frac{(2.5)^2}{2000}}} = |-5.16| = 5.16$$

$|Z_{tab}| < Z_{cat}$  we reject the ( $H_0$ ).

③ The mean yield of wheat from a district 'A' was 2.0 pounds with 10 pounds as S.D. from a sample of 100 plots in another district the mean yield was 220 pounds with S.D 12 from a sample of 150 plots assuming that S.D of yield in entire state was 11 pounds. Test whether there is any significant difference b/w the mean yield of crop in 2 districts.

$$\overline{x}_1 = 210 \quad n_1 = 100$$

$$\sigma = 11$$

$$\overline{x}_2 = 220 \quad n_2 = 150 \quad \alpha = 5\%$$

$$1) H_0 : \mu_1 = \mu_2$$

$$2) H_1 : \mu_1 \neq \mu_2$$

$$3) \alpha = 5\%$$

$$|Z_{cat}| = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \Rightarrow \frac{210 - 220}{\sqrt{\frac{(11)^2}{100} + \frac{(11)^2}{150}}}$$

$$= -7.04$$

$$|z_{cat}| = |-7.04| = 7.04$$

$$z_{tab} = 1.96$$

$z_{tab} < z_{cat}$  we reject the  $(H_0)$ .

- ④ The mean life of a sample of 10 electric bulbs was found to be 1456 hrs with S.D of 423 hrs. A second sample of 17 bulbs chosen from a different batch shows a mean life of 1280 hrs with S.D of 398 hrs. Is there any significant difference b/w means of 2 samples.

Given:

$$n_1 = 10, \bar{x}_1 = 1456, \sigma_1 = 423$$

$$n_2 = 17, \bar{x}_2 = 1280, \sigma_2 = 398$$

1.  $H_0 : \mu_1 = \mu_2$  (two tailed test)

2.  $H_1 : \mu_1 \neq \mu_2$

3.  $\alpha = 5\%$

$$|z_{tab}| = 1.96$$

$$|z_{cat}| = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{1456 - 1280}{\sqrt{\frac{(423)^2}{10} + \frac{(398)^2}{17}}} \approx 1.06$$

$$\therefore |z_{cat}| < |z_{tab}| \text{ we accept } (H_0)$$

- ⑤ Samples of students are drawn from 2 universities & their weights in kg's & means SD are calculated & given below make a large sample test. The significance of the

difference b/w means.

	Mean	S.D	Size of Samples.
Univer-A	55	10	400
Univer-B	57	5	100

Given  $n_1 = 400$ ,  $\sigma_1 = 10$ ,  $\bar{x}_1 = 55$

$n_2 = 100$ ,  $\sigma_2 = 5$ ,  $\bar{x}_2 = 57$

1)  $H_0: \mu = \mu_0$  ( $\because$  two-tailed test).

2)  $H_1: \mu \neq \mu_0$

3)  $\alpha = 5\%$ .

$$|z_{tab}| = 1.96.$$

$$|z_{cat}| = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{55 - 57}{\sqrt{\frac{10^2}{400} + \frac{5^2}{100}}} = 1.26$$

$\therefore z_{tab} > z_{cat}$  we accept the ( $H_0$ ).

\* Test of Significance of Single Proportion:

Suppose a large sample of size  $n$  is taken from a normal population to test the significant difference between the sample proportion ' $p$ ' & population proportion ' $P$ ' then

$$Z = \frac{P - p}{\sqrt{\frac{pq}{n}}}$$

$$\text{where } P = \frac{X}{n}, P + Q = 1.$$

$$Q = 1 - P.$$

① In a random sample of 125 cola drinkers 68 said they prefer thumbsup to pepsi the alternative hypothesis  $P > 0.5$ .

$$P = 0.5, n = 125, X = 68, p = \frac{X}{n} = \frac{68}{125} = 0.544$$

$$Q = 1 - P = 1 - 0.5 = 0.5$$

$$1) H_0 = P = 0.5$$

$$2) H_1 = P > 0.5$$

$$3) (\alpha) = \alpha = 5\%$$

$$|z_{\text{cat}}| = \frac{P - p}{\sqrt{\frac{pq}{n}}} = \frac{0.544 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{125}}} = 0.983.$$

$$|z_{\text{tab}}| = 1.64,$$

$|z_{\text{tab}}| > |z_{\text{cat}}|$  we accept the  $(H_0)$ .

② In a big city 325 mean out of 600 are found to be smokers thus this information support the conclusion that the majority of men in this city are smokers.

$$\text{Given } n = 600, X = 325, p = 0.5$$

$$Q = 1 - 0.5 = 0.5$$

$$p = \frac{X}{n} = \frac{325}{600} = 0.541$$

$$= \frac{325}{600} = 0.541, P > 0.5$$

$$1) (H_0) = P = 0.5 \quad (\text{Right tailed test})$$

$$2) (H_1) = P > 0.5$$

$$3) \alpha = 5\%$$

$$|z_{\text{tab}}| = 1.64.$$

$$|z_{\text{cat}}| = \frac{p - P}{\sqrt{\frac{pq}{n}}} = \frac{0.54 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{600}}} = 2.008.$$

$\therefore z_{\text{cat}} > z_{\text{tab}}$  we reject the  $(H_0)$ .

③ A manufacturer claimed that atleast 95% of the equipment which is supplied to a factory confirmed to specifications on examination of a sample of 200 species are equipment revealed that 18 are fault test is claimed at 5% level of significance.

Given  $P = 95\% = \frac{95}{100} = 0.95$

$$n = 200$$

No. of pieces confirmed to specification.

$$x = 200 - 18 \Rightarrow x = 182.$$

$$P = \frac{x}{n} = \frac{182}{200} = 0.91$$

i)  $H_0 = P = 0.91$  (right-tailed test)

ii)  $H_1 = P > 0.91$

iii) LOS  $\alpha = 5\%$ .

$$|z_{\text{cat}}| = \frac{p - P}{\sqrt{\frac{pq}{n}}} = \frac{0.91 - 0.95}{\sqrt{\frac{(0.95)(0.05)}{200}}} = 2.59.$$

$$|z_{\text{tab}}| = 1.64$$

$\therefore z_{\text{cat}} > z_{\text{tab}}$  we reject the  $(H_0)$ .

④ Among 900 peoples in a state 90 are found to be chapathi eaters. Construct 99% confidence interval for the true population.

Given  $n=900$ ,  $x=90$

$$P = \frac{x}{n} = \frac{90}{900} = \frac{1}{10} = 0.1$$

$$q = 1 - P = 1 - 0.1 = 0.9$$

Confidence interval :-

$$(P - 3\sqrt{\frac{pq}{n}}, P + 3\sqrt{\frac{pq}{n}})$$

$$= (0.1 - 3\sqrt{\frac{(0.1)(0.9)}{900}}, 0.1 + 3\sqrt{\frac{(0.1)(0.9)}{900}})$$

$$= (0.07, 0.13)$$

\* Test of significance for difference on Proportions

Proportions ~

Suppose 2 samples of size  $n_1, n_2$  are taken respectively from 2 different populations to test the significance difference b/w two samples proportions  $P_1, P_2$  then we use the formula

$$Z = \frac{P_1 - P_2}{\sqrt{PQ \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$\text{Where } P_1 = \frac{x_1}{n_1}, P_2 = \frac{x_2}{n_2}$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$Q = 1 - P$$

① Random samples of 400 men and 600 women are asked whether they would like to have a fly-over here there recidence 200 men and 325 women are in favour of the proposal. Test the hypothesis that proportions of men and women in favour of the proposal are same at 5% level of significance.

$$\text{Given } n_1 = 400, x_1 = 200$$

$$n_2 = 600, x_2 = 325$$

$$P_1 = \frac{x_1}{n_1} = \frac{200}{400} = 0.5, \quad P_2 = \frac{x_2}{n_2} = \frac{325}{600} = 0.54$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$= \frac{(400)(0.5) + (600)(0.54)}{400 + 600} = 0.52$$

$$q = 1 - P = 1 - 0.52 = 0.48$$

$$\text{i) Null hypothesis } (H_0) = P_1 = P_2$$

$$\text{ii) Alternative hypothesis } (H_1) = P_1 \neq P_2 \text{ (two-tailed test)}$$

$$\text{iii) } (\alpha) = 5\%$$

$$z_{\text{cat}} = \frac{P_1 - P_2}{\sqrt{P_2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \\ = \frac{0.5 - 0.54}{\sqrt{(0.52)(0.48) \frac{1}{400} + \frac{1}{600}}} \\ = \frac{-0.04}{\sqrt{0.00132}} = -0.24$$

$$|z_{\text{cat}}| = |-0.24| = 0.24$$

$$(z_{\text{tab}}) = 1.96$$

$$\therefore |z_{\text{tab}}| > |z_{\text{cat}}| \text{ we accept } H_0$$

② In city A 20% of a random sample of 900 school boys has a certain physical defect and another city B 18.5% of a random sample of 1600 school boys has a same defect is the difference b/w the proportions significant at 0.05 level of significance.

$$\text{Given } n_1 = 900, n_2 = 1600$$

$$x_1 = 20\% \text{ of } 900 \text{ students} = \frac{20}{100} \times 900 = 180$$

$$x_2 = 18.5\% \text{ of } 1600 \text{ students} = \frac{18.5}{100} \times 1600 = 296$$

$$P_1 = \frac{x_1}{n_1} = \frac{180}{900} = 0.2$$

$$P_2 = \frac{x_2}{n_2} = \frac{296}{1600} = 0.185$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{900(0.2) + 1600(0.185)}{900 + 1600}$$

$$= 0.1904$$

$$q = 1 - p = 1 - 0.1904$$

$$q = 0.8096$$

1) Null hypothesis ( $H_0$ ):  $P_1 = P_2$ .

2) Alternative hypothesis ( $H_1$ ):  $P_1 \neq P_2$  (two-tailed test)

3) L.O.S ( $\alpha$ ) = 5%.

$$z_{\text{cat}} = \frac{0.2 - 0.185}{\sqrt{(0.1904)(0.8096) \frac{1}{900} + \frac{1}{1600}}} = 0.916$$

$$z_{\text{tab}} = 1.96$$

$\therefore |z_{\text{tab}}| > |z_{\text{cat}}|$  we accept  $H_0$ .

③ If 2 Large populations these are 30% & 25% respectively of fair haired people is this difference slightly to be ident in samples of 1200 & 900 respectively from the two populations.

Given  $n_1 = 1200$ ,  $n_2 = 900$

$$x_1 = \frac{30}{100} = 0.3, \quad x_2 = \frac{25}{100} = 0.25$$

$P_1$  = proportion of fair haired people in 1st population i.e.

$$\frac{30}{100} = 0.3.$$

$P_2$  = proportion of fair haired people in 2nd population i.e.

$$\frac{25}{100} = 0.25$$

$$P_c = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$= \frac{(1200)(0.3) + (900)(0.25)}{1200 + 900} = 0.278.$$

$$q = 1 - p = 1 - 0.278$$

$$q = 0.722.$$

1) Null hypothesis ( $H_0$ ) =  $P_1 = P_2$

2) Alternative hypothesis ( $H_1$ ) =  $P_1 \neq P_2$

3) L.O.S. ( $\alpha$ ) = 5%.

$$|Z_{\text{cat}}| = \sqrt{\frac{0.3 - 0.25}{(0.278)(0.722) \frac{1}{1200} + \frac{1}{900}}} = 2.55.$$

$$|Z_{\text{tab}}| = 1.96.$$

$|z_{tab}| < |z_{cat}|$  we reject  $H_0$ .

- ④ In an investigation of the machine performance of 2 machines the following results are found.

Machine	No. of units inspected	no. of units defective
Machine 1	375	17
Machine 2	450	22

Given  $n_1 = 375, n_2 = 450$

$$x_1 = 17, x_2 = 22$$

$$P_1 = \frac{x_1}{n_1} = \frac{17}{375} = 0.045$$

$$P_2 = \frac{x_2}{n_2} = \frac{22}{450} = 0.048$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{375(0.045) + 450(0.048)}{375 + 450}$$

$$= 0.046$$

$$q = 1 - 0.046 = 0.953$$

1) Null hypothesis ( $H_0$ ) =  $P_1 = P_2$

2) Alternative hypothesis =  $P_1 \neq P_2$

3) L.O.S ( $\alpha$ ) = 5%.

$$|z_{cat}| = \frac{P_1 - P_2}{\sqrt{Pq(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{0.045 - 0.048}{\sqrt{(0.046)(0.953)\left(\frac{1}{375} + \frac{1}{450}\right)}}$$

$$|z_{cat}| = |-0.204| = 0.204$$

$$|z_{tab}| = 1.96$$

$|z_{tab}| > |z_{cat}|$  we accept  $H_0$

\* Small Sample :-

The size of the sample ' $n$ '  $< 30$ , then the

sample is called small sample.

They are 3 types :-

1. t-test / student 't' test

2. F-Test

3. chi-square ( $\chi^2$ ) test.

\* Degree of freedom :-

It is a number which indicates how many values of the variable may be independently (freely) chosen, it is denoted by ('v')

1. t-test / student 't' test :-

It is used for testing of hypothesis when the sample size is very small  $n < 30$ .

$\bar{x}$  = mean of a sample.

$n$  = size of the population

$\mu$  = mean of the population.

Supposed to be normal then students 't' test is defined by the test statistic.

$$t = \frac{\bar{x} - \mu}{\sqrt{s^2/n}}$$

Here  $s^2$  be the sample variance.

$$\text{If } s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Where  $s^2$  is the Unbiased estimate of the population variance  $\sigma^2$  then we use formulae:

$$\text{Test statistic } t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad (\text{one sample t-test})$$

① The average breaking strength of the student rod is specified to be 18.5 thousand pounds. To test this hypothesis 14 rods were tested. The mean and standard deviations obtained were 17.85 and 1.955 respectively. Is the result of experiment significant?

$$n=14, \bar{x}=17.85, (\text{S.D}) s=1.955$$

$$\mu = 18.5$$

$$\text{i) Null hypothesis } H_0: \mu = 18.5$$

$$\text{ii) Alternative hypothesis } H_1: \mu \neq 18.5$$

$$\text{iii) L.O.S } (\alpha) = 5\% = 0.005.$$

$$\begin{aligned} \text{Test statistic } t &= \frac{\bar{x} - \mu}{s/\sqrt{n-1}} \\ &= \frac{17.85 - 18.5}{1.955/\sqrt{14-1}} = -1.199. \end{aligned}$$

$$|z_{\text{cat}}| = 1.199.$$

At 5% level of significance for two-tailed test

$$\text{with } v=n-1$$

$$= 14-1 = 13.$$

$$z_{\text{tab}} = 2.160.$$

$\because |z_{\text{tab}}| < |z_{\text{cat}}|$  we accept  $H_0$ .

② A random sample of 6 steel beams has a mean compressive strength of 58,392 P.S.I (pounds per square inch) with a S.D. of 648 P.S.I. Use this information and the level of significance  $\alpha = 0.05$  to test whether the true average compressive strength of the street from which this sample is 58,000 P.S.I. Assume Normals.

$$n=6, \bar{H}=58,000, \bar{x}=58,392$$

$$S=648$$

$$\text{i) } N.H (H_0) : H = 58,000$$

$$\text{ii) A.H (H}_1\text{) :- } H \neq 58,000$$

$$\text{iii) L.O.S (\alpha)} = 0.05$$

$$\text{Test statistic } t = \frac{\bar{x} - H}{S/\sqrt{n-1}}$$

$$t_{\text{cat}} = 1.352$$

At 5% L.O.S for 2-tailed test with

$$H = n-1 = 6-1 = 5$$

$$t_{\text{tab}} = 2.571$$

$\therefore t_{\text{cat}} < t_{\text{tab}}$  we accept  $H_0$ .

③ A random Sample of 10 boys had following IQ's  
70, 120, 110, 101, 88, 83, 95, 98, 107, 100.

(i) Do these data support the assumption of a population mean IQ of 100

(ii) find a reasonable range in which most of the mean IQ values of sample of 10 boys lie.

$$\bar{x} = \frac{70 + 120 + 110 + 101 + 88 + 87 + 95 + 98 + 107 + 100}{10}$$

$$\bar{x} = 97.2$$

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
70	-27.2	739.84
120	-22.8	519.84
110	-12.8	163.84
101	-8.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	-2.2	4.84
98	0.8	18.336
107	9.8	96.04
100	2.8	7.84

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$= \frac{1833.6}{9} = 203.73$$

$$s = 14.27$$

$$1) N.H (H_0) = \mu = 100$$

$$2) A.H (H_1) = \mu \neq 100$$

$$3) L.O.S (\alpha) = 5\%$$

$$4) D.O.F (v) = n-1 = 10-1 = 9$$

(v)

$$t_{\text{cat}} = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

$$= \frac{97.2 - 100}{14.27/\sqrt{10}} = -0.6204.$$

$$|t_{\text{cat}}| = 0.6204.$$

$$|t_{\text{tab}}| = 2.262.$$

$|t_{\text{tab}}| > |t_{\text{cat}}|$  we accept  $H_0$

The data of support the assumption of mean IQ of 100 in the population.

(ii) The 95% confidence limits are given by

$$\left[ \bar{x} + \frac{t_{\alpha/2}}{2} \left( \frac{s}{\sqrt{n}} \right), \bar{x} - \frac{t_{\alpha/2}}{2} \left( \frac{s}{\sqrt{n}} \right) \right]$$

$$= 97.2 + \left( 2.26 \times \frac{14.27}{\sqrt{10}} \right), 97.2 - \left( 2.26 \times \frac{14.27}{\sqrt{10}} \right)$$

$$= (107.39, 87.00).$$

(iii) Producer of 'gutkha' claims that the nicotine content in its gutkha on the average is 1.83 mg. Can this claim accepted if a random sample of 8 gutkha of this type have the nicotine contents of 2, 1.7, 2.1, 1.9, 2.2, 2.1, 2, 1.6? use of 0.05 level significance.

$$\bar{x} = \frac{2+1.7+2.1+1.9+2.2+2.1+2+1.6}{8}$$

$$\bar{x} = 1.95$$

=

$x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
2.	0.05	0.0025
1.7	-0.25	0.0625
2.1	0.15	0.0225
1.9	-0.05	0.0025
2.2	0.25	0.0625
2.1	0.15	0.0225
2	0.05	0.0025
1.6	-0.35	0.1225

Students  $t$ -test for difference of means ~

Let  $\bar{x}$  and  $\bar{y}$  be the means of two independent samples of sizes  $n_1$  and  $n_2$  drawn from a normal population having means  $\mu_1$  &  $\mu_2$  then we use formulae

$$t = \frac{\bar{x} - \bar{y}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where,  $S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$

where  $r = n_1 + n_2 - 2$

If  $\bar{x}$  and  $\bar{y}$  are means and sample standard deviations  $s_1, s_2$  are given directly then we use formulae

$$t = \frac{\frac{1}{n_1} - \frac{1}{n_2}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where

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

$$\bar{x} = \frac{\sum x_i}{n_1}, \quad \bar{y} = \frac{\sum y_i}{n_2}$$

$$d.o.f = v = n_1 + n_2 - 2.$$

① Samples of 2 types of electric light bulbs are tested for length of life & following data choose.

TYPE - I	TYPE - II
Sample size $n_1 = 8$	Sample size $n_2 = 7$
Sample mean $\bar{x} = 1234$	Sample mean $\bar{y} = 1036$ hrs
Sample S.D. $s_1 = 36$ hrs	Sample S.D. $s_2 = 40$ hrs

is the difference in the mean

sufficiency that type I is superior to type-II

regarding length of life.

since, the sample sizes are small and  $\sigma_1, \sigma_2$  are not known then we use t-test. Let  $\mu_1, \mu_2$  be the 2 population means

i) N.H ( $H_0$ ):  $\mu_1 = \mu_2$

ii) A.H ( $H_1$ ):  $\mu_1 \neq \mu_2$

iii) L.D.S ( $\alpha$ ) = 0.05

Test statistic =  $t = \frac{\bar{x} - \bar{y}}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

Where  $s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$

$$S^2 = \frac{8(36)^2 + 7(40)^2 - 1848}{8+7-2}$$

$$S^2 = 1659.08$$

$$\therefore S = \sqrt{1659.08} \Rightarrow 40.73$$

$$t_{cat} = \frac{1234 - 1036}{40.73 \sqrt{\frac{1}{8} + \frac{1}{7}}} = 9.39.$$

$t_{tab}$  at 5% L.D.S. of 2-tailed test with

$$V = n_1 + n_2 - 2.$$

$$= 8 + 7 - 2 = 13 \text{ d.o.f.}$$

$$t_{tab} = 2.160.$$

$\therefore t_{cat} > t_{tab}$  we reject  $H_0$ .

(2) To examine the hypothesis that the husbands are more intelligent than the wives of an investigation took a sample of 10 couples and administered then a test which measures the IQ. The results are-

Husbands	117	105	97	105	123	109	86	78	103	107
wives	106	98	87	104	116	96	90	69	108	85

The test hypothesis with a reasonable test at the level of significance of 0.05.

Here  $n_1 = 10$ ,  $n_2 = 10$ .

$$\bar{x} = \frac{\sum x_i}{n_1} = \frac{117 + 105 + 97 + 105 + 123 + 109 + 86 + 78 + 103 + 107}{10}$$

$$\bar{x} = 103.$$

$$\bar{y} = \frac{\sum y_i}{n_2} = \frac{106 + 98 + 87 + 104 + 116 + 96 + 90 + 69 + 108 + 85}{10}$$

$$\bar{y} = 95.8$$

x	x - $\bar{x}$	$(x - \bar{x})^2$	y	$(y - \bar{y})$	$(y - \bar{y})^2$
117	14	196	106	10.2	104.04
105	2	4	98	2.2	4.84
97	-6	36	87	-8.8	77.44
105	2	4	104	8.2	67.24
123	20	400	116	20.2	408.04
109	6	36	96	-0.8	0.64
86	-17	289	90	-5.8	33.64
78	-25	625	69	-26.8	718.24
103	0	0	108	12.2	148.84
107	4	16	85	-10.8	116.64
1030		1600	958		1679.6

$$S^2 = \frac{(1600) + (1679.6)}{18} = 182.53$$

$$S = \sqrt{182.53} = 13.51$$

$$t = \frac{\bar{x} - \bar{y}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$(1) N.H (H_0) = H_1 = H_2$$

$$(2) A.H (H_1) = H_1 > H_2$$

$$3) L.D.S (\alpha) = 0.05$$

unit-4, Pg - 28/41

$$\text{Test statistic} = \frac{\bar{x} - \bar{y}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{103 - 95.8}{13.5 \sqrt{\frac{1}{10} + \frac{1}{10}}} = 1.19168$$

At  $\alpha = 0.05$  L.D.S for one tailed test (Right) for

$$V = n_1 + n_2 - 2 = 10 + 10 - 2 = 18 \text{ D.O.F.}$$

$$t_{tab} = 1.734.$$

$\therefore t_{cat} > t_{tab}$  we accept  $H_0$ .

- ③ Two horses A and B were tested according to the time (in sec) to run a particular track with the following results.

Horse A 28 30 32 33 33 29 34.

Horse B 29 30 30 24 27 29 30

Test whether the two horses have the same running capacity.

Here  $n_1 = 7$ ,  $n_2 = 6$ .

$$\bar{x} = \frac{\sum x_i}{n_1} = \frac{28+30+32+33+33+29+34}{7} = 31.283.$$

$$\bar{y} = \frac{\sum y_i}{n_2} = \frac{29+30+30+24+27+29}{6} = 28.166.$$

$x$	$x - \bar{x}$	$(x - \bar{x})^2$	$y$	$(y - \bar{y})$	$(y - \bar{y})^2$
28	-3.28	10.75	29	0.884	0.781
30	-1.28	1.651	30	1.834	3.363
32	0.715	0.5112	30	1.834	3.363
33	0.1.715	2.941	24	-4.166	17.355
33	1.715	2.941	27	-1.16	1.345
29	-2.285	5.221	29	0.834	0.695
34	2.715	7.371			
		31.38			26.90

$$S^2 = \frac{31.38 + 26.90}{7+6-2} = 5.298$$

$$S = \sqrt{5.298} = 2.301$$

(i)  $H_0: \mu_1 = \mu_2$

(ii)  $H_1: \mu_1 \neq \mu_2$

iii) L.O.S ( $\alpha$ ) :- 0.05

$$\begin{aligned} \text{Test statistic} &= \frac{\bar{x} - \bar{y}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\ &= \frac{31.285 - 28.166}{2.301 \sqrt{\frac{1}{7} + \frac{1}{6}}} = 2.436 \end{aligned}$$

At  $\alpha = 0.05$  level of significance for two tailed

test for  $v = n_1 + n_2 - 2$

$$= 7+6-2 = 11 \text{ d.o.f.}$$

$$t_{tab} = 2.201$$

$\therefore t_{cat} > t_{tab}$  we reject  $H_0$

Unit 4, Pg - 30(4)

④ Two independent samples of 8 & 7 items respectively had the following:

sample I 11, 11, 13, 11, 15, 9, 12, 14

sample II 9, 11, 10, 13, 9, 8, 10

Is the difference b/w the means of Samples significant.

Here,  $n_1 = 8$ ,  $n_2 = 7$

$$\bar{x} = \frac{\sum x_i}{n_1} = \frac{11+11+13+11+15+9+12+14}{8} = 12$$

$$\bar{y} = \frac{\sum y_i}{n_2} = \frac{9+11+10+13+9+8+10}{7} = 10$$

x	$(x - \bar{x})$	$(x - \bar{x})^2$	y	$(y - \bar{y})$	$(y - \bar{y})^2$
11	-1	1	9	-1	1
11	-1	1	11	1	1
13	1	1	10	0	0
11	-1	1	13	3	9
15	3	9	9	-1	1
9	3	9	8	-2	4
12	0	0	10	0	0
14	2	4			
		26			16

② \* F-Test :- (Equality of variances)

Let 2 independent random samples of sizes  $n_1, n_2$  be drawn from 2 normal populations, to test the hypothesis that the 2 population variances  $\sigma_1^2$  and  $\sigma_2^2$  are equal.

Let the Null hypothesis be ( $H_0$ ) :  $\sigma_1^2 = \sigma_2^2$  against the Alternative hypothesis be ( $H_1$ ) :  $\sigma_1^2 \neq \sigma_2^2$

The estimates of  $\sigma_1^2$  &  $\sigma_2^2$  are given by,

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{\sum (x_i - \bar{x})^2}{n_1 - 1} \text{ and}$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{\sum (y_i - \bar{y})^2}{n_2 - 1}$$

where  $s_1^2$  and  $s_2^2$  are the variance of 2 samples we use the test statistic for t-test is

$$F = \frac{s_1^2}{s_2^2} \quad [\text{if } s_1^2 > s_2^2]$$

(or)

$$F = \frac{s_2^2}{s_1^2} \quad [\text{if } s_2^2 > s_1^2]$$

with

$$\begin{cases} v_1 = n_1 - 1 \\ v_2 = n_2 - 1 \end{cases}$$

degree of freedom.

① pumpkins were grown under 2 experimental conditions. two random samples of 11 and 9. Show the sample standard deviations of their weights 0.8 & 0.5 respectively assuming that the weight distribution is normal test hypothesis that the two variances are equal.

$$n_1 = 11, n_2 = 9 \text{ (Sample size)}$$

$$S_1^2 = 0.8, S_2^2 = 0.5 \text{ (Sample variance)}$$

$$S_1^2 = \frac{n_1 S_1^2}{n_1 - 1} = \frac{11(0.8)^2}{11 - 1} = 0.704$$

$$S_2^2 = \frac{n_2 S_2^2}{n_2 - 1} = \frac{9(0.5)^2}{9 - 1} = 0.281$$

$$\text{clearly } S_1^2 > S_2^2$$

$$1) \text{N.H. } (H_0) : \sigma_1^2 = \sigma_2^2$$

$$2) \text{A.H. } (H_1) : \sigma_1^2 \neq \sigma_2^2$$

$$3) \text{L.O.S. } (\alpha) : 0.05$$

$$\text{Test statistic } F = \frac{S_1^2}{S_2^2} = \frac{0.704}{0.281}$$

$$F_{\text{cat}} = 1.57$$

$$F_{\text{tab}} = 3.35$$

At 5% L.O.S. for F-test with

$$v_1 = n_1 - 1 = 11 - 1 = 10$$

$$v_2 = n_2 - 1 = 9 - 1 = 8$$

$\because F_{\text{tab}} > F_{\text{cat}}$  we accept  $H_0$

② In two independent samples of sizes 8 & 10, the sum of squares of deviations of the sample values from the respective sample mean were 84.4 & 102.6 test whether the difference of the variables of the population is significant (or) not.

$$n_1 = 8, n_2 = 10 \text{ and } \sum (x_i - \bar{x})^2 = 84.4$$

$$S_1^2 = 102.6, S_2^2 = 0.05$$

$$\sum (x_i - \bar{x})^2 = 84.4$$

$$\sum (y_i - \bar{y})^2 = 102.6$$

$$S_1^2 = \frac{\sum (x_i - \bar{x})^2}{n_1 - 1} = \frac{84.4}{8 - 1} = 12.05$$

$$S_2^2 = \frac{\sum (y_i - \bar{y})^2}{n_2 - 1} = \frac{102.6}{10 - 1} = 11.4$$

$$1) N.H (H_0) = \sigma_1^2 = \sigma_2^2$$

$$2) A.H (H_1) = \sigma_1^2 \neq \sigma_2^2$$

$$3) L.O.S (\alpha) = 0.05$$

$$\text{Test statistics } F = \frac{S_1^2}{S_2^2} = \frac{12.05}{11.4}$$

$$F_{cat} = 105$$

5% L.O.S of F-test with  $v_1 = n_1 - 1 = 7$

$$n_2 = n_2 - 1 = 9$$

$$F_{tab} = 3.29$$

$\therefore F_{cat} > F_{tab}$  we accept  $(H_0)$ .

③ In one sample of 8 observations from a normal population, the sum of the squares of deviations of the sample values from the sample mean is 84.4 and in another sample of 10 observations it was 102.6. Test at 5% level whether the population have the same variance.

$$n_1 = 8, n_2 = 10 \text{ and } S_1^2 = 84.4, S_2^2 = 102.6$$

$$\sum (x_i - \bar{x})^2 = 84.4$$

$$\sum (y_i - \bar{y})^2 = 102.6$$

$$1) H_0: \sigma_1^2 = \sigma_2^2$$

$$2) H_1: \sigma_1^2 \neq \sigma_2^2$$

$$3) L.O.S (\alpha) = 5\%$$

$$S_1^2 = \frac{1}{n_1-1} \sum (x_i - \bar{x})^2 = \frac{84.4}{7} = 12.057$$

$$S_2^2 = \frac{1}{n_2-1} \sum (y_i - \bar{y})^2 = \frac{102.6}{9} = 11.4$$

$$F = \frac{S_1^2}{S_2^2} = \frac{12.057}{11.4} = 1.057$$

Degrees of freedom are given by

$$v_1 = n_1 - 1 = 8 - 1 = 7$$

$$v_2 = n_2 - 1 = 10 - 1 = 9$$

$$T = 3.29 \text{ at } 5\%$$

We accept the Null hypothesis ( $H_0$ )

Q) The nicotine contents in milligrams in two samples of tobacco were found to be as follows:-

Sample A    24    27    26    21    25

Sample B    27    30    28    31    22    36

Can it be said that the two samples

have come from the same normal population?

Given  $n_1=5$ ,  $n_2=6$ .

Calculation for mean's and S.D's of samples

$x$	$x-\bar{x}$	$(x-\bar{x})^2$	$y$	$(y-\bar{y})$	$(y-\bar{y})^2$
24	0.6	0.36	27	-2	4
27	2.4	5.76	30	+1	1
26	1.4	1.96	28	-1	1
21	3.6	12.96	31	+2	4
25	0.4	0.16	22	-7	49
			36	+7	49
123	21.2	174			108

$$\bar{x} = \frac{\sum x_i}{n_1} = \frac{123}{5} = 24.6$$

$$\bar{y} = \frac{\sum y_i}{n_2} = \frac{174}{6} = 29.$$

$$\sum (x_i - \bar{x})^2 = 21.2, \quad \sum (y_i - \bar{y})^2 = 108$$

$$S_1^2 = \frac{\sum (x_i - \bar{x})^2}{n_1 - 1} = \frac{21.2}{4} = 5.3$$

$$S_2^2 = \frac{\sum (y_i - \bar{y})^2}{n_2 - 1} = \frac{108}{5} = 21.6$$

small samples:-

i) F-Test :-

Test statistic for the difference of two variances is given by

$$\text{Null hypothesis } H_0 = \sigma_1^2 = \sigma_2^2$$

$$F = \frac{s_2^2}{s_1^2} = \frac{21.6}{5.9} = 4.075$$

Tabulated value of F for  $(n_2-1, n_1-1)$

$$= (5, 4) \text{ d.f at } 5\% \text{ LOS is } 2.26$$

As calculated value is greater than tabulated value  
∴ we accept the  $H_0$

ii) Student's t-test :-

The 2 means are equal i.e.  $\mu_1 = \mu_2$

Given  $\bar{x} = 24.6, \bar{y} = 29$ .

$$\sum (x_i - \bar{x})^2 = 21.2, \sum (y_i - \bar{y})^2 = 108$$

$$S^2 = \frac{1}{n_1 + n_2 - 2} [\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2]$$

$$S^2 = \frac{21.2 + 108}{5+6-2} (21.2 + 108) = 14.35$$

$$S = 3.78$$

The test statistic is  $t = \frac{\bar{x} - \bar{y}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$$t = \frac{24.6 - 29}{3.78 \sqrt{\frac{1}{5} + \frac{1}{6}}} = \frac{-4.4}{3.78} = -1.192$$

Table values of t for 9 dof at 5% LOS is 2.26

⇒ from ① & ② we conclude that the 2 samples come from same normal populations.

### ③ \* CHI - SQUARE ( $\chi^2$ ) Test (Test of fit)

If  $O_i$  ( $i=1, 2, 3, \dots, n$ ) is a set of observed (Experimental) frequencies and  $E_i$  ( $i=1, 2, \dots, n$ ) frequencies

Then  $\chi^2$  is defined as

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Where the  $(n-1)$  degree of freedom  $\chi^2$  is used

to test whether differences b/w observed & expected frequencies are significant.

(i) The no. of auto mobile accidents for week in a

certain community are as follows 12, 8, 20, 2, 14, 10, 15

6, 9 and 4 All these frequencies in agreement with the belief that accident conditions were the same during this 10 week period. Expected frequencies of accidents per week  $= \frac{100}{10} = 10$ .

(i) N.H ( $H_0$ ) : The accidents were the same during 10 week period.

(ii) A.H ( $H_1$ ) : The accidents are different during 10 week period

$$\text{Ans}(\alpha) = 0.05$$

Explain the null hypothesis and alternative hypothesis

Difference between 10 and 1 week

O.F ( $O_i$ )	E.F ( $E_i$ )	$O_i - E_i$	$\frac{(O_i - E_i)^2}{E_i}$
12	10	2	0.4
8	10	-2	0.4
20	10	10	10.0
2	10	-8	6.4
14	10	4	1.6
10	10	0	0.0
15	10	5	2.5
6	10	-4	1.6
9	10	-1	0.1
4	10	-6	3.6
100	100		26.6

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = 26.6$$

At 5% L.O.S for  $\chi^2$ -test with  $v=n-1=10-1=9$

$$\chi^2_{tab} = 16.919$$

$\therefore \chi^2_{cat} > \chi^2_{tab}$  we reject  $(H_0)$ .

- ② A die is thrown 264 times with the following results show that the die is unbiased given that ( $\chi^2_{0.05} = 1.07$  for 5 d.o.f).

No appeared on die	1	2	3	4	5	6
Frequency	40	32	28	58	54	52

The expected frequency of each of the

numbers 1, 2, 3, 4, 5, 6 is  $\frac{264}{6} = 44$ .

i) N.H ( $H_0$ ) = The die is unbiased.

ii) A.H ( $H_1$ ) = The die is biased.

O.F ( $O_i$ )	E.F ( $E_i$ )	$O_i - E_i$	$\frac{(O_i - E_i)^2}{E_i}$
40	44	-4	0.36
32	44	-12	3.27
28	44	-16	5.81
58	44	14	4.45
54	44	10	2.27
52	44	8	1.45
264	264		17.6362

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = 17.6362$$

At 5% L.O.S. for  $\chi^2$  test with  $N = n - 1 = 6 - 1 = 5$

$$\chi_{tab} = 11.070$$

$\therefore \chi_{cat} > \chi_{tab}$  we reject  $H_0$ .

③ A pair of dies are thrown 360 and the frequency of each sum is indicated below:

Sum	2	3	4	5	6	7	8	9	10	11	12
frequency	8	24	35	37	44	65	51	42	26	14	14

Would you say that the dies are fair

on the bases of the  $\chi^2$ -test at 0.05 L.O.S.

(i)  $H_0$  : The dies are fair

(ii)  $H_1$  : The dies are not fair.

(iii) LOS ( $\alpha$ ) : 0.05

The probability of getting sum 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

$x = x_i$	2	3	4	5	6	7	8	9	10	11	12
$P(x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

sum	$O_i$ (O.F)	$E \cdot F (E_i)$ $= 360 \times P(x_i)$	$(O_i - E_i)$	$\frac{(O_i - E_i)^2}{E_i}$
2	8	$\frac{1}{36} \times 360 = 10$	$8 - 10 = -2$	0.4
3	24	20	$24 - 20 = 4$	0.8
4	35	30	$35 - 30 = 5$	0.83
5	37	40	$37 - 40 = -3$	0.225
6	44	50	$44 - 50 = -6$	0.72
7	65	60	$65 - 60 = 5$	0.41
8	51	50	$51 - 50 = 1$	0.02
9	42	40	$42 - 40 = 2$	0.1
10	26	30	$26 - 30 = -4$	0.53
11	14	20	$14 - 20 = -6$	1.8
12	14	10	$14 - 10 = 4$	1.6

At 5% LOS for  $\chi^2$  test

with  $v = n - 1 = 10$

$$\chi_{tab} = 19.67$$

$\because \chi_{tab} > \chi_{cat}$  we accept  $H_0$ .