Test of Hypothesis

1. **Define Sample.**
   Solution: A Sample is a part of the statistical population (i.e) it is a subset which is collected to draw an inference about the population.
2. **Define Sample size.**
   Solution: The number of individuals in a sample is called the sample size
3. **Define Null hypotheses and Alternative hypothesis.**
   Solution: For applying the test of significance, we first set up of a hypothesis, a definite statement about the population parameter, such a hypothesis is usually called as null hypothesis and it is denoted by H_0.
   Any hypothesis which is complementary to the null hypothesis is called an alternative hypothesis and it is denoted by H_1.
4. **A random sample of 200 tins of coconut oil gave an average weight of 4.95 kgs with SD of 0.21 kg. Do we accept the hypothesis of net weight 5 kgs per tin at 1% level ? Explain. (L6)**
   Solution:
   Sample size n=200
   Sample mean $\bar{x}$=4.95kg
   Sample SD s=0.21kg
   Population mean μ=5kg.
   The sample is a large sample and so apply z-test.
   H_0 : μ=5kg
   H_1 : μ≠5kg
   The test statistic is $z=\frac{\bar{x}-\mu}{s/\sqrt{n}}$
   $$=\frac{4.95-5}{0.21/200}=\frac{-.05\times\sqrt{200}}{0.21}=-3.37$$
   ∴ $|z|$=3.37
   $At\ 1\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ is\ 2.58.$
   Conclusion: H_0 is rejected at 1% *level* since calculated value of $|z|$ is greater than the table value of z.Therefore the net weight tin is not equal to 5 kg.
5. **A sample of 900 items has mean 3.4 and SD 2.61. Test weather the sample be regarded as drawn from a population with mean 3.25 at 5% level of significance? (L4)**
   Solution:
   Sample size n=900

Sample mean$\bar{x}$=3.4

Sample SD s=2.61

Population mean μ=3.25

The sample is a large sample and so apply z-test.

H_0 : μ=3.25

H_1 : μ≠3.25

The test statistic is z=$\frac{\bar{x}-\mu}{s/\sqrt{n}}$

$$=\frac{3.4-3.25}{2.61/\sqrt{900}}=\frac{0.15}{2.61/30}=1.72$$

∴ $|z|$=1.72

$At\ 1\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ is\ 2.58.$

Conclusion: H_0 is accepted at 1% *level* since calculated value of $|z|$ is less than the table value of z.Therefore H_0 is accepted.

6. **A Sample of 400 male students is found to have a mean height of 171.38 cms. Can it be reasonable regarded as a sample from a large population with mean height 171.17 cms and standard deviation 3.30 cms? Justify? (L6)**

Solution:

Sample size n=400

Sample mean$\bar{x}$=171.38cm

Population SDσ=3.30cm

Population mean μ=171.17cm

The sample is a large sample and so apply z-test.

H_0 : μ=171.17cm

H_1 : μ≠171.17cm

The test statistic is z=$\frac{\bar{x}-\mu}{s/\sqrt{n}}$

$$=\frac{171.38-171.17}{3.30/\sqrt{400}}=\frac{0.21\times 20}{3.30}=1.27$$

∴ $|z|$=1.72

$At\ 5\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ is\ 1.96.$

Conclusion: H_0 is accepted at 5% *level* since calculated value of $|z|$ is less than the table value of z.Therefore H_0 is accepted andμ=171.17cm.

7. **The mean of two samples of 1000 and 2000 numbers are respectively 67.5 and 68 inches. Can they be regarded as draws from the same population with SD 2.5 inches? Justify? (L6)**

Solution:

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

n_1=1000, n_2=2000

Population SDσ=2.5

The two given samples are large samples.

$H_0 : \mu_1=\mu_2$

$H_1 : \mu_1 \neq \mu_2$

The test statistic is z=$\frac{\bar{x}_1-\bar{x}_2}{\sigma/\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ = $\frac{67.5-68}{2.5/\sqrt{\frac{1}{1000}+\frac{1}{2000}}}$ =-6.25

∴ $|z|$=6.25

$At\ 1\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ is\ 2.58.$

Conclusion: H_0 is rejected at 1% *level* since calculated value of $|z|$ is greater than the table value of z.

∴H_0 is rejected at 1% *level* of significance and so the two samples cannot be regarded as belonging to the same population.

8. The random samples of sizes 400 and 500 have mean 10.9 and 11.5 respectively. Can the samples be regarded as drawn from the same population with variance 25? Justify? (L6)

Solution:

$\bar{x}_1$=10.9, $\bar{x}_2$=11.5

n_1=400, n_2=500

σ^2=25

The two given samples are large samples.

$H_0 : \mu_1=\mu_2$

$H_1 : \mu_1 \neq \mu_2$

The test statistic is z=$\frac{\bar{x}_1-\bar{x}_2}{\sigma/\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ = $\frac{10.9-11.5}{5/\sqrt{\frac{1}{400}+\frac{1}{500}}}$ =-2.38

∴ $|z|$=2.38

$At\ 1\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ is\ 2.58.$

Conclusion: H_0 is accepted at 1% *level* since calculated value of $|z|$ is less than the table value of z.

Therefore the samples come from the population with variance 25.

9. A sample of 26 bulbs given a mean life of 990 hours with a SD of 20 hours. The manufactures claims that the mean life of bulbs is 1000 hours. Is the sample not upto the standard? Justify? (L6)

Solution:

Sample size n=26$< 30(small\ sample)$

Sample mean$\bar{x}$=990

Sample SD s=20

Population mean μ=1000

Degrees of freedom=n-1=26-1=25

Here we know $\bar{x}$,μ,SD and n.Therefore, we use student's 't' test.

H_0 : The sample is upto the standard.

H_1 : The sample is not upto the standard.

The test statistic is t=$\frac{\bar{x}-\mu}{s/\sqrt{n-1}}$

$$=\frac{990-1000}{20/\sqrt{25}}=-2.5$$

∴ $|t|$=2.5 (i.e) Calculated t=2.5

$At\ 5\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ at\ 25d.f\ is\ 2.06$

Conclusion:$H_0$is rejected as calculated value is greater than the tabulated value. ∴The sample is not upto the standard.

10.In one sample of 8 observations the sum of the squares of deviations of the sample values from the sample mean was 84.4 and it the other sample of 10 observations it was 102.6. Test whether this difference is significant at 5% level? (L4)

Solution:

n_1=8, n_2=10

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

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

H_0 :$S_1^2=S_2^2$

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

(i.e) calculated F=1.057

Tabulated value of F for(7,9) degrees of freedom is 3.29.

Calculated value F<Tabulated value F

∴We accept the null hypothesis.

11.A sample of size 13 gave an estimates population variance of 3.0, while another sample of size 15 gave an estimate of 2.5. Could both samples be from populations with the same variance. Justify? (L6)

Solution:

n_1=13, n_2=15

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

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

$H_0 : S_1^2 = S_2^2$.The two samples have come from populations with same variance.

∴The test statistic is

$F=\frac{S_1^2}{S_2^2} = \frac{(Greater\ variance)}{(Smaller\ variance)} = \frac{3.0}{2.5}=1.2$

(i.e) calculated F=1.2

Tabulated value of F for(12,14) degrees of freedom is 2.53

Calculated value F<Tabulated value F

∴We accept the null hypothesis H_0

$(i.e)$ Both samples have come from the populations with the same variance.

12.Write the test procedure of Chi-square test? (L5)

Solution:

(i) Write down the null hypothesis

(ii) Write down the alternative hypothesis.

(iii) Calculate the theoretical frequencies for the contingency.

(iv) Calculate $\aleph^2=\sum\frac{(O-E)^2}{E}$

(v) Write down the number of degress of freedom.

(vi) Write the conclusion on the hypothesis by comparing the calculated values of $\aleph^2$with table value of $\aleph^2$

13.Write the uses of $\aleph^2$ – test? (L1)

Solution:

(i) It is used to test the goodness of a distribution.

(ii) It is used to test the significance of the difference between the observed frequencies in a sample and the expected frequencies,obtained from the theoretical distribution.

(iii)It is also used to test the independence of attributes.

(iv)In case of small samples(where the population standard deviation is not known) $\aleph^2$ statistic is used to test whether a specified value can be the population variance σ^2.

14.A machine is designed to produce insulation washers for electrical devices of average thickness of 0.025cm. A random sample of 10 washers was found to have a thickness of 0.024cm with a S.D of 0.002 cm. Test the significance of the deviation value of t for 9 degrees of freedom at 5% $level$ is 2.262. (L4)

Solution:

Sample size n=10$< 30(small\ sample)$

Sample mean$\bar{x}$=0.024cm

Sample SD s=0.002cm

Population mean μ=0.025cm

Degrees of freedom=n-1=10-1=9

Here we know $\bar{x}$,μ,SD and n.Therefore, we use student's 't' test.

H_0 : The difference between $\bar{x}$ and μ is not significant

The test statistic is t=$\frac{\bar{x}-\mu}{s/\sqrt{n-1}} = -1.5$

∴ $|t|$=1.5 (i.e) Calculated t=1.5

$At\ 5\%\ level\ of\ significance\ the\ tabulated\ value\ of\ z\ at\ 9d.f\ is\ 2.06$

Conclusion:H_0 is accepted as calculated value is less than the tabulated value.

PART-B

1. **Find student's t, for the following variate value in a sample of eight -4, -2,-2,0,2,2,3,3 taking the mean of the universe to be zero.(L1)**

Solution:

Number of samples=8

∴n=8

Mean of universe is zero

∴ μ=0

$\bar{X}$=average value of X

=$\frac{(-4)+(-2)+(-2)+0+2+2+3+3}{8}$

=0.25

To calculate S,we have the formula

S = $\sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$

Hypothesis: There is no significant difference between sample mean and population mean

X	$X-\bar{X}$	$(X-\bar{X})^2$
-4	-4.25	18.06
-2	-2.25	5.06
-2	-2.25	5.06
0	-0.25	0.06
2	1.75	3.06
2	1.75	3.06
3	2.75	7.56
3	2.75	7.56

$$\sum (X-\bar{X})^2 = 49.98$$

S = $\sqrt{\frac{\sum(X-\bar{X})^2}{n-1}} = \sqrt{\frac{49,48}{8-1}}$ =$\sqrt{5.497}$ =2.658

t=$\frac{\bar{X}-\mu}{S/\sqrt{n-1}}$=$\frac{0.25-0}{2.658/\sqrt{7}}$ =0.248

Table value=2.26

∴calculated value <tabulated value

∴ Hypothesis is accepted and so there is no significant difference between sample mean and population mean.

2. **Ten students are selected at random in a university and their heights are measured in inches as 64,65,65,67,67,69,69,70,72 and 72.Using these data, Discuss the suggestion that the mean height of the students in the university is 66.(At 5% level of significance the value of t for 9 d.f is 2.262).(L2)**

Solution:

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

S = $\sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$

$\bar{X}$=average value of X

=$\frac{64+65+65+67+67+69+69+70+72+72}{10}$ =68

Hypothesis: There is no significant difference between sample height and population height.

X	$X-\bar{X}$	$(X-\bar{X})^2$
64	-4	16
65	-3	9
65	-3	9
67	-1	1
67	-1	1
69	1	1
69	1	1
70	2	4
72	4	16
72	4	16

$$\sum (X-\bar{X})^2 = 74$$

$S=\sqrt{\frac{74}{10-1}}=\sqrt{\frac{74}{9}}$ =2.867

Here $\bar{X}$=68, $\mu = 66, \mathrm{n} = 10$

$t = \frac{68-66}{2.867/\sqrt{10}}$=2.205

Table value=2.26

∴calculated value <tabulated value,therefore Hypothesis is accepted and the height of population group can be taken as 66.

3. **A fertilizer mixing machine is set to give 12kg of nitrate for every quintal bag of fertilizer.Ten 100kg bags are examined.The percentages of nitrate are as follows 11,14,13,12,13,12,13,14,11,12.Is there reason to belive that the machines is defective? (value of t for 9 d.f is 2.262). Justify? (L6)**

Solution:

Hypothesis: There is no significant difference between sample percentage and population percentage.

Here n=10

$$\mu = 12$$

$\bar{X}$=average value of X

$=\frac{11+14+13+12+13+12+13+14+11+12}{10}$ =12.5

X	$X-\bar{X}$	$(X-\bar{X})^2$
11	-1.5	2.25
14	1.5	2.25
13	0.5	0.25
12	-0.5	0.25
13	0.5	0.25
12	-0.5	0.25
13	0.5	0.25
14	1.5	2.25
11	-1.5	2.25
12	-0.5	0.25

$$\sum (X-\bar{X})^2 = 10.5$$

To calculate S,we have the formula

$S = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}=\sqrt{\frac{10.5}{10-1}}$

S=1.08

$t = \frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{12.5-12}{1.08} \times 3$

t=1.389

Table value=2.26

∴calculated value <tabulated value

∴ Hypothesis is accepted and the machine cannot be believed to be defective.

4. Two random samples drawn from two normal populations are given below.Test whether the two populations have the same variances (L4)

Samples I	**20**	**16**	**26**	**27**	**23**	**22**	**18**	**24**	**25**	**19**		
Samples II	**17**	**23**	**32**	**25**	**22**	**24**	**28**	**6**	**31**	**20**	**33**	**27**

Solution:

Hypothesis: There is no significant difference between variances of the two samples.

By Formula

$F = \frac{S_1^2}{S_2^2}$ if $S_1^2 > S_2^2$

$= \frac{S_2^2}{S_1^2}$ if $S_2^2 > S_1^2$

$$where\ S_1^2 = \sum \frac{(X_1-\overline{X_1})^2}{n_1-1}$$

$$S_2^2 = \sum \frac{(X_2-\overline{X_2})^2}{n_2-1}$$

Here $n_1 = 10, n_2$=12

Calculating the averages of two samples we get,

$$\overline{X_1} = 22, \overline{X_2} = 24$$

X_1	$X_1-\overline{X_1}$	$(X_1-\overline{X_1})^2$	X_2	$X_2-\overline{X_2}$	$(X_2-\overline{X_2})^2$
20	-2	4	17	-7	49
16	-6	36	23	-1	1
26	4	16	32	8	64
27	5	25	25	1	1
23	1	1	22	-2	4
22	0	0	24	0	0
18	-4	16	28	4	16
24	2	4	6	-18	324
25	3	9	31	7	49
19	-3	9	33	9	81
			20	-4	16

			27	3	9

$\sum(X_1 - \overline{X_1})^2 = 120 \sum(X_2 - \overline{X_2})^2 = 614$

$$S_1^2 = \sum \frac{(X_1 - \overline{X_1})^2}{n_1 - 1} = \frac{120}{9} = 13.33$$

$$S_2^2 = \sum \frac{(X_2 - \overline{X_2})^2}{n_2 - 1} = \frac{614}{11} = 55.81 \qquad \Rightarrow S_2^2 > S_2^2$$

$\therefore F = \frac{S_2^2}{S_1^2} = \frac{55.81}{13.33} = 4.18$

Degrees of freedom $\gamma_1 = 12 - 1 = 11, \quad \gamma_2 = 10 - 1 = 9$

Table value=3.10

Calculated value $>$ Table value

$\therefore$ Hypothesis is rejected.

$\therefore$ There is significant difference between the variance.

5. In two groups of ten children each increases in weight due to two different diets in the same period were in pounds.

8	**5**	**7**	**8**	**3**	**2**	**7**	**6**	**5**	**7**
3	**7**	**5**	**6**	**5**	**4**	**4**	**5**	**3**	**6**

Find whether the variance are significantly different . (L1)

Solution :

Ho : there is no significant Diffenence between the variance of the two samples

F = $\frac{S_1^2}{S_2^2}$ If $S_1^2 > S_2^2 = \frac{S_2^2}{S_1^2}$ $if \quad S_2^2 > S_1^2$

Where $S_1^2 = \frac{\sum(X_1 - \overline{X_1})^2}{n_1 - 1}$ Here n_1=10 n_2= 10

$S_2^2 = \frac{\sum(X_2 - \overline{X_2})^2}{n_2 - 1}$ $\overleftarrow{X_1}$ =5.8$\overleftarrow{X_1}$=4.8

X_1	$X_1 - \overleftarrow{X_1}$	$(X_1 - \overleftarrow{X_1})^2$	X_2	$X_2 - \overleftarrow{X_2}$	$(X_2 - \overleftarrow{X_2})^2$
3	2.2	4.84	3	-1.8	3.24
5	-0.8	0.64	7	2.2	4.84
7	1.2	1.44	5	0.2	0.04
8	2.2	4.84	6	1.2	1.44
3	-2.8	7.84	5	0.2	0.04
2	-3.8	14.44	4	-0.8	0.64
7	1.2	1.44	4	-0.8	0.64
6	0.2	0.04	5	0.2	0.04
5	-0.8	1.64	3	-1.8	3.24
7	1.2	1.44	6	1.2	1.44

		$\sum(X_1 - \overleftarrow{X_1})^2$ =37.6			$\sum(X_2 - \overleftarrow{X_2})^2$ =15.6

$S_1^2 = \frac{\sum(X_1 - \overline{X_1})^2}{n_1 - 1} = \frac{37.6}{9}$=4.18

$S_2^2 = \frac{\sum(X_2 - \overline{X_2})^2}{n_2 - 1} = \frac{15.6}{9}$=1.73

$$Here \quad S_1^2 > S_2^2$$

F = $\frac{S_1^2}{S_2^2} = \frac{4.18}{1.73}$ =2.42

$$Here$$

v_1=10-1=9 v_2=10-1=9

Degrees of freedom =9.9

Table value for the Degrees of freedom 9.9 at 5% level =3.23

Calculated value =2.42 $< Table\ value$

∴ H_0 =Accepted

There is no significant deffeerence between the variance .

6. **The nicotine contents in milligrams in two samples of tobacco were found to be as follows.**

Samples A	**24**	**27**	**26**	**21**	**25**	
Samples B	**27**	**30**	**28**	**31**	**22**	**36**

Can it be said that the two samples have same variance.Justify?(L6)

Solution :

H_0 = There is no significant deffeerence between the variance of the two 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

$\bar{X} = \frac{\sum X}{n} = \frac{123}{5}$ =24.6

$\bar{Y} = \frac{\sum Y}{n} = \frac{174}{6}$ =29

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

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

F= $\frac{S_2^2}{S_1^2}$=$\frac{21.6}{5.3}$ =4.07

Calculated value =4.07

Table value of F for $(5,4)$ d,f at 5% level is 6.26

$\therefore$ calculated value calculated value $<$ Table value.

$\therefore$ We accept H_0 ie ; The variance are equal .

7. Two random samples were drawn from two normal populations and their values are

A	**66**	**67**	**75**	**76**	**82**	**84**	**88**	**90**	**92**		
B	**64**	**66**	**74**	**78**	**82**	**85**	**87**	**92**	**93**	**95**	**97**

Test whether the two populations have the same variance at 5% level of Significance. (L4)

Solution :

There is no significant difference between the variance of the sample .

X	X-$\bar{X}$	$(X-\bar{X})^2$	Y	Y-$\bar{Y}$	$(Y-\bar{Y})^2$
66	-14	196	64	-19	361
67	-13	169	66	-17	289
75	-5	25	74	-9	81
76	-4	16	78	-5	25
82	2	4	82	-1	1
84	4	16	85	2	4
88	8	64	87	4	16
90	10	100	92	9	81
92	12	144	93	10	100
			95	12	144
			97	14	196
720	0	734	913	0	1298

S_1^2 =$\frac{\sum(X-\bar{X})^2}{n_1-1}$ =$\frac{734}{8}$=91. S_2^2 =$\frac{\sum(Y-\bar{Y})^2}{n_2-1}$ = $\frac{1298}{10}$=129.8

$S_2^2 > S_1^2$

F= $\frac{S_2^2}{S_1^2}$=$\frac{129.8}{91.75}$ =1.41

Degree of freedom is $(10,8)$

Table value of F =3.34 AT 5% LEVEL

$\therefore$ calculated value $<$ Table value.

$\therefore$ WE Accepted H_0.

There is no significant difference between the variance of the two population .