SUBJECT NAME: ENGINEERING MATHEMATICS II

(COMMON TO BIO GROUPS)

SUBJECT CODE: SMT1106

COURSE MATERIAL

UNIT- V THEORY OF SAMPLING AND TESTING OF HYPOTHESIS

Population: The group of individuals, under study is called population.

Sample: A finite subset of statistical individuals in a population is called Sample.

Sample size: The number of individuals in a sample is called the Sample size.

Parameters and Statistics: The statistical constants of the population are referred as Parameters and the statistical constants of the Sample are referred to as Statistics.

Standard Error: The standard deviation of sampling distribution of a statistic is known as its standard error and is denoted by (S.E)

Test of Significance: It enables us to decide on the basis of the sample results if the deviation between the observed sample statistic and the hypothetical parameter value is significant or the deviation between two sample statistics is significant.

Null Hypothesis: A definite statement about the population parameter which is usually a hypothesis of no-difference and is denoted by Ho.

Alternative Hypothesis: Any hypothesis which is complementary to the null hypothesis is called an Alternative Hypothesis and is denoted by H_1.

if $\mu = \mu_0$ is the null hypothesis H_0 then, the alternate hypothesis H_1 could be $\mu > \mu_0$ (Right tail) or $\mu < \mu_0$ (Left tail) or $\mu \neq \mu_0$ (Two tail test)

Errors in Sampling: Type I and Type II errors.

Type I error: Rejection of H_0, when it is true.

Type II error: Acceptance of H_0, when it is false.

Critical region: A region corresponding to a statistic "t" in the sample space S which leads to the rejection of H_o is called Critical region or Rejection region.

Acceptance Region: Those regions which lead to the acceptance of H_o are called Acceptance Region.

Level of Significance: The probability α that a random value of the statistic "t" belongs to the critical region is known as the level of significance.

Types of samples: Small sample and Large sample. A sample is said to be a small sample if the size is less than or equal to 30 otherwise it is a large sample.

Large Sample

Z test for mean

Test of significance for single Mean

$Z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$, where $\bar{x}$ the sample mean, μ is is the population mean, σ is the population standard deviation and n is the sample size.

Test of significance for difference of mean

$Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)}}$, where $\bar{x}_1$ is the first sample mean, $\bar{x}_2$ is the second sample mean, n_1 is the first sample size, n_2 is the second sample size, s_1^2 is the first sample variance and s_2^2 is the second sample variance.

Confidence Limits

The values of $\bar{x} \pm 1.96 \dfrac{\sigma}{\sqrt{n}}$ are called 95% confidence limits for the mean of the population corresponding to the given sample. The values of $\bar{x} \pm 2.58 \dfrac{\sigma}{\sqrt{n}}$ are called 99% confidence limits for the mean of the population corresponding to the given sample.

Z test for proportions

Test of significance for single proportion

$Z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$, where P is the population proportion, Q = 1 – P, p is the sample proportion and n is the sample size.

Test of significance for difference of proportion

$Z = \dfrac{p_1 - p_2}{\sqrt{PQ\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$, where p_1 is the first sample proportion, p_2 is the second sample proportion, n_1 is the first sample size, n_2 is the second sample size, $P = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$ and Q = 1 – P

Small Sample

t -Test of significance for single Mean

$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$, where $\bar{x}$ the sample mean, μ is is the population mean, s is the sample standard deviation and n is the sample size.

If the mean and standard deviation are not given, then the following formulae are used to calculate

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

Degrees of freedom is n - 1

Confidence Limits

Let $\bar{x}$ be the sample mean and n be the sample size. Let s be the sample standard deviation. Then the 95 % level confidence limits are given by $\bar{x} \pm t_{0.05} \frac{s}{\sqrt{n-1}}$. The 99 % level confidence limits are given by $\bar{x} \pm t_{0.01} \frac{s}{\sqrt{n-1}}$.

Test of significance for difference of mean

$Z = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$, where $\bar{x}_1$ is the first sample mean, $\bar{x}_2$ is the second sample mean, n_1 is the first sample size, n_2 is the second sample size, $s^2 = \frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2}$.

Degrees of freedom is $n_1 + n_2 - 2$

F test

F= $\frac{Greater\ variance}{Smaller\ variance}$ i.e, $F = \frac{S_1{}^2}{S_2{}^2}$ if $S_1{}^2 > S_2{}^2$ (OR) $F = \frac{S_2{}^2}{S_1{}^2}$ if $S_2{}^2 > S_1{}^2$

If the sample variances s_1^2 and s_2^2 are given, then the following formula can be used to calculate S_1^2 and S_2^2 :

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1}, S_2^2 = \frac{n_2 s_2^2}{n_2 - 1}$$

If the sample variances s_1^2 and s_2^2 are not given and the set of observations for both samples are given then the following formula can be used to calculate S_1^2 and S_2^2

$S_1^2 = \frac{\sum (x-\bar{x})^2}{n_1 - 1}$, $S_2^2 = \frac{\sum (y-\bar{y})^2}{n_2 - 1}$, where n_1 is the first sample size, n_2 is the second sample size, $\bar{x}$ is the first sample mean and $\bar{y}$ is the second sample mean.

χ^2 test

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Where O is the observed frequency and E is the expected frequency.

Calculation of expected frequencies in testing independence of attributes

Expected Frequency = (Row total * Column Total) / Grand total

Explanation for the above with two classes is given below

Observed Frequencies

			Total
	a	c	a+c
	b	d	b+d
Total	a+b	c+d	a+b+c+d = N

Expected Frequencies

			Total
	E(a) = $\frac{(a+c)(a+b)}{N}$	E(c) = $\frac{(a+c)(c+d)}{N}$	a+c
	E(b) = $\frac{(b+d)(a+b)}{N}$	E(d) = $\frac{(c+d)(b+d)}{N}$	b+d
Total	a+b	c+d	a+b+c+d = N

Problems

1. **A company manufacturing electric light bulbs claims that the average life of its bulbs is 1600 hours. The average life and standard deviation of a random sample of 100 such bulbs were 1570 hours and 120 hours respectively. Test the claim of the company at 5% level of significance.**

Solution:

Null Hypothesis H_0: $\mu = 1600$. There is no significant difference between sample mean and population mean

Alternative Hypothesis H_1: $\mu \neq 1600$. There is a significant difference between sample mean and population mean.

The statistic test is $Z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

$$Z = \frac{1570 - 1600}{\dfrac{120}{\sqrt{100}}} = -2.5$$

$$|z| = 2.5$$

Calculated value z = 2.5

Tabulated value of z at 5% level of significance for a two tail test is 1.96

Calculated value >Tabulated value, H_0 is rejected.

We cannot accept the claim of the company.

2. **The breaking strength of ropes produced by a manufacturer has mean 1800N and standard deviation 100N. By introducing a new technique in the manufacturing process it is claimed that the breaking strength has increased. To test this claim a sample of 50 ropes is tested and it is found that the breaking strength is 1850N. Can we support the claim at 1% level of significance?**

Solution:

Null Hypothesis H_0: μ = 1800 N

Alternative Hypothesis H_1: μ > 1800 N (one tailed test)

n = 50, $\bar{x}$ = 1850 μ =1800 σ = 100

The statistic test is $Z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

$$Z = \frac{1850 - 1800}{\dfrac{100}{\sqrt{50}}} = 3.54$$

Calculated value z = 3.54

Tabulated value of z at 5% level of significance for a one tail test is 2.33

Calculated value >Tabulated value, H_0 is rejected.

The difference is significant and so we support the claim of the manufacturer.

3. **Measurements of the weights of a random sample of 200 ball bearings made by a certain machine during one week showed a mean of 0.824N and a standard deviation of 0.042N. Find the 95% and 99% confidence limits for the mean weight of all the ball bearings.**

Solution:

The 95% confidence limits are $\bar{x} \pm 1.96 \frac{s}{\sqrt{n}}$

n = 200, $\bar{x}$ = 0.824 s = 0.042

$$\bar{x} \pm 1.96 \frac{s}{\sqrt{n}} = 0.824 \pm (1.96)\left(\frac{0.042}{\sqrt{200}}\right) = 0.824 \pm 0.0058$$

The 95% confidence interval is (0.8182, 0.8298)

The 99% confidence limits are $\bar{x} \pm 2.58 \frac{s}{\sqrt{n}}$

$$\bar{x} \pm 2.58 \frac{s}{\sqrt{n}} = 0.824 \pm (2.58)\left(\frac{0.042}{\sqrt{200}}\right) = 0.824 \pm 0.0077$$

The 99% confidence interval is (0.8163, 0.8317)

4. **In a survey of buying habits, 400 women shoppers are chosen at random in supermarket A. Their average weekly food expenditure is Rs.250 with standard deviation Rs.40. For 400 women shoppers chosen at random in supermarket B, the average weekly food expenditure is Rs.220 with standard deviation is Rs.55. Test at 1% level of significance whether the average weekly food expenditure of the populations of shoppers are equal.**

Solution:

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

$n_1 = 400 \quad n_2 = 400 \quad \bar{x} = 250 \quad \bar{y} = 220 \quad s_1 = 40 \quad s_2 = 55$

$$s = \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)} = \sqrt{\frac{40^2}{400} + \frac{55^2}{400}} = 3.4$$

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} = \frac{250 - 220}{3.4} = 8.82$$

Calculated value z = 8.82

Tabulated value of z at 1% level of significance for a two tailed test is 2.56

Calculated value >Tabulated value, H_0 is rejected.

The difference in the weekly food expenditure is significantly different.

5. **A random sample of 500 pineapples was taken from a large consignment and 65 were found to be bad. Test whether the proportion of bad ones is not significantly different from 0.1 at 1% level of significance**

Solution:

Null Hypothesis H_0: P = 0.1There is no significant difference between sample and population proportion.

Alternative Hypothesis H_1: $P \neq 0.1$ There is a significant difference between sample and population proportion.

The statistic test is $Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$

$$p = \frac{65}{500} = 0.13$$

P = 0.1 Q= 1 – P = 1 – 0.1= 0.9

$$Z = \frac{0.13-0.1}{\sqrt{\frac{(0.1)(0.9)}{500}}} = 2.238$$

Calculated value z = 2.238

Tabulated value of z at 5% level of significance for a two tail test is 1.96

Calculated value >Tabulated value, H_0 is rejected.

The proportion of bad ones is significantly different from 0.1

6. **In a sample of 1000 people, 540 were rice eaters and the rest were wheat eaters. Can we assume that the proportion of rice eaters is more than 50% at 1% level of significance.**

Solution:

H_0: P = 0.5

H_1: P > 0.5 (One tailed test)

P = 0.5 Q= 1 – P = 1 – 0.5 = 0.5

$$p = \frac{540}{1000} = 0.54$$

The statistic test is $Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$

$$Z = \frac{0.54-0.5}{\sqrt{\frac{(0.5)(0.5)}{1000}}} = 2.532$$

Calculated value z = 2.532

Tabulated value of z at 5% level of significance for a one tail test is 2.33

Calculated value >Tabulated value, H_0 is rejected.

The rice eaters are more than 50% of the population.

7. **In a random sample of 900 votes, 55% are favored the Democratic candidate for the post of the President. Test the hypothesis that the Democratic candidate has more chances of winning the President post.**

Solution:

H_0: P = 0.5

H_1: P > 0.5 (Right tailed test)

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

$p = \frac{55}{100} = 0.55$

The statistic test is $Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$

$$Z = \frac{0.55 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{900}}} = 3$$

Calculated value z = 3
Tabulated value of z at 5% level of significance for a one tail test is 2.33
Calculated value >Tabulated value, H_0 is rejected.
The Democratic candidate is having more chances to win the President Post.

8. In a random sample of 1000 persons from town A, 400 are found to be consumers of wheat. In a sample of 800 from town B, 400 are found to be consumers of wheat. Do these data reveal a significant difference between town A and town B so far as the proportion of wheat consumers is concerned?

Solution:
H_0: Two towns do not differ much as far as the proportion of wheat consumption. $P_1 = P_2$
H_1: $P_1 \neq P_2$

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

$$p_1 = \frac{400}{1000} = 0.4 \qquad p_2 = \frac{400}{800} = 0.5$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{1000(0.4) + 800(0.5)}{1000 + 800} = 0.444$$

$$Q = 1 - P = 1 - 0.444 = 0.556$$

$$Z = \frac{0.4 - 0.5}{\sqrt{(0.444)(0.556)\left(\frac{1}{1000} + \frac{1}{800}\right)}} = \frac{0.1}{0.024} = 4.17$$

Calculated value z = 4.17
Tabulated value of z at 5% level of significance for a two tail test is 1.96
Calculated value >Tabulated value, H_0 is rejected.
Hence the data reveal a significant difference between town A and town B so far as the proportion of wheat consumers is concerned.

9. In the past, a machine has produced washers having a thickness of 0.050 inch. To determine whether the machine is in proper working order, a sample of 10 washers is chosen, for which the mean thickness is 0.053 inch and the standard

deviation is 0.003 inch. Test the hypothesis that the machine is in proper working order, using 5% and 1% level of significance.

Solution:
H_0: μ = 0.050
H_1: μ $\neq$ 0.050 (two tailed test)
$n = 10 \qquad \bar{x} = 0.053 \qquad s = 0.003 \qquad \mu = 0.050$

The statistic test is $t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n-1}}} = \dfrac{0.053 - 0.050}{\dfrac{0.003}{\sqrt{10-1}}} = 3.00$

Calculated value t = 3.00
Degree of freedom = n – 1 =10 – 1 = 9
At 5% LOS:
Tabulated value of t at 5% level of significance with 9 degrees of freedom for a two tailed test is 2.26
Calculated value >Tabulated value, H_0 is rejected.
The Machine is not in proper working order at 5% level of significance
Tabulated value of t at 1% level of significance with 9 degrees of freedom for a two tailed test is 3.25
Calculated value < Tabulated value, H_0 is accepted.
The Machine is in proper working order at 1% level of significance.

10. The specifications for a certain kind of ribbon call for a mean breaking strength of 180 pounds. If five pieces of the ribbon (randomly selected from the different rolls) have a mean breaking strength of 169.5 pounds with a standard deviation of 5.7 pounds. Test the null hypothesis $\mu = 180$ pounds against the alternative hypothesis $\mu < 180$ pounds at the 0.01 level of significance. Assume that the population distribution is normal.

Solution:
H_0: $\mu = 180$
H_1: $\mu < 180$ (left tailed test)
$n = 5 \qquad \bar{x} = 169.5 \qquad s = 5.7 \qquad \mu = 180$

The statistic test is $t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n-1}}} = \dfrac{169.5 - 180}{\dfrac{5.7}{\sqrt{5-1}}} = -3.68$

Calculated value t = 3.68
Degree of freedom = n – 1 =5 – 1 = 4
Tabulated value of t at 1% level of significance with 4 degrees of freedom for a one tail test is 3.747.
Calculated value > Tabulated value, H_0 is accepted.
Hence the mean breaking strength can be taken as 180 pounds.

11. Ten individuals are chosen at random from a normal population and their heights are found to be 63,63,66,67,68,69,70,70,71,71 inches. Test the hypothesis that the mean height is greater than 66 inches at 5% level of significance

Solution:

H_0: μ = 66

H_1: μ > 66 (one tailed test)

$$\bar{x} = \frac{\sum x}{n} = \frac{678}{10} = 67.8$$

X	63	63	66	67	68	69	70	70	71	71	Total
$(x-\bar{x})$	- 4.8	- 4.8	- 1.8	- 0.8	0.2	1.2	2.2	2.2	3.2	3.2	
$(x-\bar{x})^2$	23.04	23.04	3.24	0.64	0.04	1.44	4.84	4.84	10.24	10.24	81.6

$$s^2 = \frac{\sum (x-\bar{x})^2}{n} = \frac{81.6}{10} = 8.16$$

s = 2.857

The statistic test is $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}} = \frac{67.8-66}{\frac{2.857}{\sqrt{9}}} = 1.89$

Calculated value t = 1.89

Degree of freedom = n – 1 =10 – 1 = 9

Tabulated value of t at 5% level of significance with 9 degrees of freedom for a one tail test is 1.833

Calculated value >Tabulated value, H_0 is rejected. Accepted H_1

The Mean is significantly higher than 66 inches.

12. Two independent samples of size 8 and 7 items had the following values

Sample I	**9**	**11**	**13**	**11**	**15**	**9**	**12**	**14**
Sample II	**10**	**12**	**10**	**14**	**9**	**8**	**10**	

Test if the difference between the mean is significant

Solution:

H_0: $\mu_1 = \mu_2$ There is no significant difference between means

H_1: $\mu_1 \neq \mu_2$ There is a significant difference between means

$$\bar{x} = \frac{\sum x}{n} = \frac{94}{8} = 11.75 \qquad \bar{y} = \frac{\sum y}{n} = \frac{73}{7} = 10.43$$

x	(x-$\bar{x}$)	$(x-\bar{x})^2$	y	(y-$\bar{y}$)	$(y-\bar{y})^2$
9	- 2.75	7.56	10	- 0.43	0.185
11	- 0.75	0.56	12	1.57	2.465
13	1.25	1.56	10	- 0.43	0.185
11	- 0.75	0.56	14	3.47	12.041
15	3.25	10.56	9	- 1.43	2.045
9	- 2.75	7.56	8	- 2.43	5.905
12	0.25	0.06	10	- 0.43	0.185
14	2.25	5.06			
94		33.48	73		23.011

$$s^2 = \frac{\sum(x-\bar{x})^2 + \sum(y-\bar{y})^2}{n_1+n_2-2} = \frac{33.48+23.011}{8+7-2} = 4.35$$

s = 2.086

The statistic test is $t = \dfrac{\bar{x}-\bar{y}}{s\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} = \dfrac{11.75-10.43}{2.086\sqrt{\dfrac{1}{8}+\dfrac{1}{7}}} = 1.22$

Calculated value t = 1.22

Degree of freedom = n_1+n_2-2 = 8 + 7 – 2 = 13

Tabulated value of t at 5% level of significance with 13 degrees of freedom for a two tail tets is 2.16

Calculated value < Tabulated value, H_0 is accepted

There is no significant difference between means.

13. The IQ of 16 students from one area of a city showed a mean of 107 with the standard deviation 10, while the IQ of 14 students from another area showed a mean of 112 with standard deviation 8. Is there a significant difference between the IQ's of the two groups at 1% and 5% level of significance?

Solution:

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

$n_1 = 16$ $\quad n_2 = 14$ $\quad s_1 = 10$ $\quad s_2 = 8$ $\quad \bar{x} = 107$ $\quad \bar{y} = 112$

$$s^2 = \frac{n_1 s_1^{\ 2} + n_2 s_2^{\ 2}}{n_1+n_2-2} = \frac{16(10)^2 + 14(8)^2}{16+14-2} = \frac{2496}{28} = 89.143$$

s = 9.44

The statistic test

$$t = \frac{\bar{x}-\bar{y}}{s\sqrt{\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}} = \frac{107-112}{9.44\sqrt{\dfrac{1}{16}+\dfrac{1}{14}}} = -1.45$$

Calculated value t = 1.45

Degree of freedom = $n_1 + n_2 - 2$ = 16 + 14 – 2 = 28
At 5% LOS:
Tabulated value of t at 5% level of significance with 28 degree of freedom for a two tail test is 2.05
Calculated value < Tabulated value, H_0 is accepted
There is no significant difference in the IQ level of the two groups.

At 1% LOS:
Tabulated value of t at 1% level of significance with 28 degree of freedom is 2.76
Calculated value < Tabulated value, H_0 is accepted.
There is no significant difference in the IQ level of the two groups.

14. A random sample of 10 parts from machine A has a sample standard deviation of 0.014 and another sample of 15 parts from machine B has a sample standard deviation of 0.08. Test the hypothesis that the samples are from a population with same variance.

Solution:

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

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

$n_1 = 10 \quad n_2 = 15 \quad s_1 = 0.014 \quad s_2 = 0.08$

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{10 \times (0.014)^2}{10 - 1} = 0.0002$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{15 \times (0.08)^2}{15 - 1} = 0.006$$

$$F = \frac{S_2^2}{S_1^2} = \frac{0.006}{0.0002} = 30$$

Calculated value F = 30
Tabulated Value of F at 5% level of significant with (14, 9) degrees of freedom is 3.03
Calculated value > Tabulated value, H_0 is rejected
There is a significant difference in the variances of two populations.

15. Two random samples drawn from two normal populations are

Sample I	**20**	**16**	**26**	**27**	**23**	**22**	**18**	**24**	**25**	**19**		
Sample II	**27**	**33**	**42**	**35**	**32**	**34**	**38**	**28**	**41**	**43**	**30**	**37**

Obtain the estimates of the variances of the population and test whether the two populations have the same variance.

Solution:

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

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

$$\bar{x} = \frac{\sum x}{n} = \frac{220}{10} = 22$$

$$\bar{y} = \frac{\sum y}{n} = \frac{420}{12} = 35$$

x	(x-$\bar{x}$)	$(x-\bar{x})^2$	y	(y-$\bar{y}$)	$(y-\bar{y})^2$
20	- 2	4	27	- 8	64
16	- 6	36	33	- 2	4
26	4	16	42	7	49
27	5	25	35	0	0
23	1	1	32	- 3	9
22	0	0	34	- 1	1
18	- 4	16	38	3	9
24	2	4	28	- 7	49
25	3	9	41	6	36
19	- 3	9	43	8	64
			30	- 5	25
			37	2	4
		120			314

$n_1 = 10 \qquad n_2 = 12$

$$S_1^{\,2} = \frac{\sum (x-\bar{x})^2}{n_1 - 1} = \frac{120}{9} = 13.33$$

$$S_2^{\,2} = \frac{\sum (y-\bar{y})^2}{n_2 - 1} = \frac{314}{11} = 28.54$$

$$F = \frac{S_2^2}{S_1^2} = \frac{28.54}{13.33} = 2.14$$

Calculated value F = 2.14

Tabulated Value of F at 5% level of significance with (11, 9) degrees of freedom is 3.1

Calculated value < Tabulated value, H_0 is accepted

There is no significant difference between variances.

16. In one sample of 8 observations the sum of squares of deviations of the sample values from the sample mean was 84.4 and in another sample of 10 observations it was 102.6. Test whether this difference is significant at 5% level.

Solution:

H$_0$: $\sigma_1^2 = \sigma_2^2$

H$_1$: $\sigma_1^2 \neq \sigma_2^2$

$n_1 = 8 \quad n_2 = 10 \quad \sum(x-\bar{x})^2 = 84.4 \quad \sum(y-\bar{y})^2 = 102.6$

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

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

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

Calculated value F = 1.057

Tabulated Value of F at 5% level of significance with (7, 9) degrees of freedom is 3.29

Calculated value < Tabulated value, H$_0$ is accepted

There is no significant difference between variances.

17. The mean life of a sample of 9 bulbs was observed to be 1309 hrs with standard deviation 420 hrs. A second sample of 16 bulbs chosen from a different batch showed a mean life of 1205 hrs with a standard deviation 390 hrs. Test at 5% level whether both the samples come from the same normal population.

Solution:

Both t-test and F-test has to be done to check whether they have come from the same population. Fist F-test is done and the followed by t-test.

F-test:

H$_0$: $\sigma_1^2 = \sigma_2^2$

H$_1$: $\sigma_1^2 \neq \sigma_2^2$

$n_1 = 9 \quad n_2 = 16 \quad s_1 = 420 \quad s_2 = 390 \quad \bar{x} = 1309 \quad \bar{y} = 1205$

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{9 \times (420)^2}{9-1} = 198450$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{16 \times (390)^2}{16-1} = 162240$$

$$F = \frac{S_1^2}{S_2^2} = \frac{198450}{162240} = 1.223$$

Calculated value F = 1.223

Tabulated Value of F at 5% level of significant with (15, 8) degree of freedom is 3.22

Calculated value < Tabulated value, H$_0$ is accepted .

t-test:

H$_0$: $\mu_1 = \mu_2$

H$_1$: $\mu_1 \neq \mu_2$

$$s^2 = \frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2} = \frac{9(420)^2 + 16(390)^2}{9 + 16 - 2} = \frac{4021200}{23} = 174834.7826$$

s = 418.13

The statistic test

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{1309 - 1205}{418.13\sqrt{\frac{1}{9} + \frac{1}{16}}} = \frac{104}{174.22} = 0.596$$

Calculated value t = 0.596

Degree of freedom = $n_1 + n_2 - 2$ = 9 + 16 – 2 = 23

Tabulated value of t at 5% level of significance with 23 degree of freedom is 2.069

Calculated value < Tabulated value, H$_0$ is accepted

Since in both F-Test and t-Test we have accepted the null hypothesis, we conclude that the samples have come from the same normal populations.

18. A dice is tossed 120 times with the following results:

No. turned up	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
Frequency	**30**	**25**	**18**	**10**	**22**	**15**	**120**

Test the hypothesis that the dice is unbiased.

Solution:

Null Hypothesis H$_0$: The dice is an unbiased one.

Alternative Hypothesis H$_1$: The dice is biased

O	E	O – E	$(O-E)^2$	$\left[\frac{(O-E)^2}{E}\right]$
30	20	10	100	5.00
25	20	5	25	1.25
18	20	- 2	4	0.20
10	20	-10	100	5.00
22	20	2	4	0.20
15	20	- 5	25	1.25
				12.90

Calculated $\chi^2 = \left[\frac{(O-E)^2}{E}\right]$ = 12.90

Degree of freedom = n – 1 = 6 – 1 = 5

Calculated value of χ^2 at 5% level of significance with 5 degree of freedom is 11.07

Tabulated value = 11.07

Tabulated value > calculated value, H_0 is rejected.
The dice are biased.

19. Genetic theory states that children having one parent of blood type M and other of blood type N will always be one of the three types M, MN, N and that the ratios of these types will be 1:2:1. A report states that out of 300 children having one M parent and one N parent, 30% were found to be of type M, 45% of type MN and remainder type N. Test the hypothesis using χ^2 test.

Solution:
H_0: There is no significant difference between the theoretical ratio and observed ratio.
H_1: There is no significant difference between the theoretical ratio and observed ratio.

If theoretical ratio is true the 300 children should be distributed as follows:

$$Type\ \ M\ = \frac{1}{4} \times 300 = 75$$

$$Type\ \ MN\ = \frac{2}{4} \times 300 = 150$$

$$Type\ \ N\ = \frac{1}{4} \times 300 = 75$$

Observed frequencies:

$$Type\ \ M\ = \frac{30}{100} \times 300 = 90$$

$$Type\ \ MN\ = \frac{45}{100} \times 300 = 135$$

$$Type\ \ N\ = \frac{25}{100} \times 300 = 75$$

Type	observed	Expected	$O-E$	$(O-E)^2$	$\left[\frac{(O-E)^2}{E}\right]$
M	90	75	15	225	3
MN	135	150	- 15	225	1.5
N	75	75	0	0	0
Total					4.5

Calculated $\chi^2 = \left[\frac{(O-E)^2}{E}\right] = 4.5$
Degree of freedom = n – 1 = 3 – 1 = 2
Calculated value of χ^2 at 5% level of significance with 2 degree of freedom is 5.99
Tabulated value = 5.99
Calculated value < Tabulated value, H_0 is accepted
There is no significant difference between the theoretical ratio and observed ratio.

20. A certain drug was administered to 456 males, out of a total 720 in a certain locality, to test its efficacy against typhoid. To incidence of typhoid is shown below. Find out the effectiveness of the drug against the disease. (The table value of $\chi^2$for 1 degree of freedom at 5% level of significance is 3.84)

	Infection	No Infection	Total
Administering the drug	**144**	**312**	**456**
Without administering the drug	**192**	**72**	**264**
Total	**336**	**384**	**720**

Solution:

Null Hypothesis H_0: The drug is independent.

Alternative Hypothesis H_1: The drug is not independent

The expected frequencies are

$\frac{336 \times 456}{720} = 212.8$ ≈213	$\frac{384 \times 456}{720} = 243.2$ ≈243	456
$\frac{336 \times 264}{720} = 123.2$ ≈123	$\frac{384 \times 264}{720} = 140.8$ ≈141	264
336	384	720

O	E	O – E	$(O-E)^2$	$\left[\frac{(O-E)^2}{E}\right]$
144	213	- 69	4761	22.35
192	123	69	4761	38.71
312	243	69	4761	19.59
72	141	- 69	4761	33.77
				114.42

Calculated χ^2= $\left[\frac{(O-E)^2}{E}\right]$= 114.42

Degree of freedom = (r – 1) (c – 1)= (2-1)(2-1) = 1

Tabulated value of χ^2 at 5% level of significance with 1 degree of freedom is 3.841

Tabulated value = 3.841

Calculated value >Tabulated value, H_0 is rejected.

Therefore, the drug is definitely effective in controlling the typhoid.

21. A brand Manager is concerned that her brand's share may be unevenly distributed throughout the country. In a survey in which the country was divided into four geographical regions, a random sampling of 100 consumers in each region was surveyed, with the following results:

	Region				
	NE	**NW**	**SE**	**SW**	**TOTAL**
Purchased the brand	**40**	**55**	**45**	**50**	**190**
Did not purchase	**60**	**45**	**55**	**50**	**210**

Using χ^2 test, find out if the brand is unevenly distributed throughout the country.

Solution:

H_0: There is no significant difference between the observed and expected frequencies

H_1: There is a significant difference between the observed and expected frequencies

The expected frequencies are :

	Region				
	NE	NW	SE	SW	TOTAL
Purchased the brand	$\frac{190 \times 100}{400} \approx 47$	$\frac{190 \times 100}{400} \approx 48$	$\frac{190 \times 100}{400} \approx 47$	$\frac{190 \times 100}{400} \approx 48$	190
Did not purchase	$\frac{210 \times 100}{400} \approx 53$	$\frac{210 \times 100}{400} \approx 52$	$\frac{210 \times 100}{400} \approx 53$	$\frac{210 \times 100}{400} \approx 52$	210
	100	100	100	100	400

O	E	O – E	$(O-E)^2$	$\left[\frac{(O-E)^2}{E}\right]$
40	47	- 7	49	1.04
55	48	7	49	1.02
45	47	- 2	4	0.085
50	48	2	4	0.083
60	53	7	49	0.924
45	52	-7	49	0.942
55	53	2	4	0.075
50	52	- 2	4	0.076
				4.245

Calculated χ^2= $\left[\frac{(O-E)^2}{E}\right]$= 4.245

Degree of freedom = (r – 1)(c – 1)= (2-1)(4-1) = 3

Tabulated value of χ^2 at 5% level of significance with 3 degree of freedom is 7.815

Tabulated value = 7.815

Calculated value < Tabulated value, H_0 is accepted

There is no significant difference between the observed and expected frequencies.