P & S

UNIT - IV

[part A : 1 Mark]

1. write down the Applications of the F-distribution?

A. The F-distribution used in finance to test whether the variances of stock returns are equal across two or more portfolios. It is also used in Engineering to test the effectiveness of different manufacturing process by Comparing the variances of the outcomes.

2. what is that Degree of freedom?

A. The number of independent variables which make up the Statistic is known as the degree of freedom and it is denoted by ν.

e.g :- In a set of data of n observations, if k is the number of independent Constraints then $\nu = n-k$

3. write the applications of chi-square distribution?

A. * To Test the goodness of fit.
* To Test the independence of attributes.
* To Test the homogeneity of independent estimation of the population variance.

4. write the uses of the t-test?

A. * To test the significance of the sample mean, when population variance is not given.

* To test the significance of the mean of the sample
* To test the significance of the difference between two sample means or to compare two samples.

5. Write the formula for chi-square test?

A. The chi-square formula, is,

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

Where O_i = observed value (actual value)

E_i = expected value.

[part B : 3 Mark]

6. producer of gutkha claims that the nicotine content in his gutkha on the average is 1.83 mg. Can this claim accepted if a random sample of 8 gutkha items of this type have the nicotine contents of 2.0, 1.7, 2.1, 1.9, 2.2, 2.1, 2.0, 1.6 mg? Use 0.05 level of significance

Sol: Given

n = Sample size = 8

$$\bar{x} = \text{Sample mean} = \frac{2.0 + 1.7 + 2.1 + 1.9 + 2.2 + 2.1 + 2.0 + 1.6}{8}$$

$$= 1.95$$

μ = population mean = 1.83

x	$x-\bar{x}$	$(x-\bar{x})^2$
2.0	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	0.05	0.0025
1.6	-0.35	0.1225
15.6		0.3

$$s^2 = \frac{1}{(n-1)}\sum_i (x_i-\bar{x})^2 = \frac{1}{7}(0.3) = 0.042$$

$$s = \sqrt{0.042} = 0.21$$

s = Sample S.D = 0.21

ν = degree of freedom = $n-1 = 7$

(i) NULL Hypothesis : $\mu = 1.83$, i.e the nicotine Content in Gutkha of the average is 1.83 mg.

(ii) Alternative Hypothesis : $\mu \neq 0$

(iii) level of Significance : $\alpha = 0.05$ (5%) assumed

(iv) Test Statistic :

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n-1}} = \frac{1.5-0}{6.8/\sqrt{7}} = \frac{1.5\times\sqrt{7}}{6.8}$$

$$= 0.583$$

The critical value of t at 5% level of Significance for 7 degree of freedom of two-tailed test is

$$t_{\alpha/2} = t_{0.025} = 2.36$$

(v) Conclusion

As $|t| < 2.36$, the NH H_0 is accepted.

$\therefore$ There is no improvement by taking the Coaching.

7. A random Sample of Size 25 from a normal population has the mean $\bar{x} = 47.5$ and S.D $= 8.4$. Does this information tend to Support to refuse the claim that the mean of the population is $\mu = 42.5$?

Sol. n = Sample Size = 25

$\bar{x}$ = Sample mean = 47.5

μ = population mean = 42.5

s = Sample S.D = 8.4

Test statistic is

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{47.5 - 42.5}{8.4/\sqrt{25}} = \frac{5 \times 5}{8.4}$$

$$= \frac{25}{8.4} = 2.98$$

V = degree of freedom = $n-1$ = 24

The critical value of t at 5% L.O.S with 24 d.f is

$$t_{\alpha/2} = t_{0.025} = 2.064$$

Since, $|t| > 2.064$, the information tend to refuse the client that the mean of the population is $\mu = 42.5$

8. The average breaking strength of the steel rods is specified to be 18.5 thousand pounds. To test this sample of 14 rods were tested. The mean and standard deviations obtained were 17.85 and 1.955 respectively. Is the result of experiment significance.?

Sol:- n = Sample Size = 14

$\bar{x}$ = Sample mean = 17.85

μ = population mean = 18.5

s = S.D of sample = 1.955

Degree of freedom = $\nu = n-1 = 14-1 = 13$.

(i) Null Hypothesis : $\mu = 18.5$ i.e the result of experiment is not significant.

(ii) Alternative Hypothesis : $\mu \neq 18.5$

(iii) Level of Significance : $\alpha = 0.05$ (5%) assumed

(iv) Test Statistic : $t = \dfrac{\bar{x}-\mu}{s/\sqrt{n-1}} = \dfrac{17.85 - 18.5}{1.955/\sqrt{13}}$

$= \dfrac{-0.65 \times \sqrt{13}}{1.955} = -1.198 \Rightarrow |t| = 1.198$

The Critical Value of $t_{\alpha/2}$ at 5% level of significance with degree of freedom 13 is 2.16.

(v) Conclusion:

As $|t| < 2.16$, the Null Hypothesis H_0 is accepted.

$\therefore$ The result of experiment is not significant.

9. The means of two random samples of sizes 9 and 7 are 196.42 and 198.82 respectively. The sum of the squares of the deviations from the mean is 26.94 and 18.73 respectively. The sample the be considered to have been drawn from the same normal population.

Sol: $n_1 = 9$, $n_2 = 7$

$\bar{x} = 196.42$, $\bar{y} = 198.82$ and

$\Sigma(x_i - \bar{x})^2 = 26.94$, $\Sigma(y_i - \bar{y})^2 = 18.73$

$$\therefore S^2 = \frac{\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2}{n_1 + n_2 - 2} = \frac{26.94 + 18.73}{9 + 7 - 2}$$

$S^2 = 3.26$

$\therefore \boxed{S = 1.81}$

(i) NH (H_0): The two samples are drawn from the sample population. i.e $\mu_1 = \mu_2$

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

(iii) Level of significance: $\alpha = 0.05$

(iv) The test statistic is, $t = \dfrac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \dfrac{196.42 - 198.82}{1.81\sqrt{\frac{1}{9} + \frac{1}{7}}}$

$t = -2.63$

The critical value of t at 5% L.o.S with degree of freedom is $t_{\alpha/2}^{14} = 2.15 < 2.63$

We reject the Null Hypothesis and conclude that the two samples are not drawn from the same population

10. The number of automobile accidents per week in a certain Community are as follows: 12, 8, 20, 2, 14, 10, 15, 6, 9, 4. Are these frequencies in agreement with the belief that accident Conditions were the Same during this 10 week period.

Sol: let

$O_1 = 12 \qquad O_6 = 10$

$O_2 = 8 \qquad O_7 = 15$

$O_3 = 20 \qquad O_8 = 6$

$O_4 = 2 \qquad O_9 = 9$

$O_5 = 14 \qquad O_{10} = 4$ be the given observed frequency

Then, the expected frequencies of accidents each week is

$$E_i = \frac{\sum_{i=1}^{10} O_i}{10} = \frac{100}{10} = 10 \quad \forall\ i = 1, 2, 3 \ldots 10$$

$\therefore E_1 = 10,\ E_2 = 10,\ E_3 = 10,\ E_4 = 10,\ E_5 = 10,\ E_6 = 10,\ E_7 = 10,$

$E_8 = 10,\ E_9 = 10,\ E_{10} = 10$

O_i	E_i	$(O_i - E_i)$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
12	10	2	4	0.4
8	10	-2	4	0.4
20	10	10	100	10.0
2	10	-8	64	6.4
14	10	4	16	1.6
10	10	0	0	0
15	10	5	25	2.5
6	10	-4	16	1.6
9	10	-1	1	0.1
4	10	6	36	3.6

$O_i = 100$, $E_i = 100$, $(O_i - E_i)^2 / E_i = 26.6$

(i) NH (H_0): The accident Conditions were the Same during the 10 week period.

(ii) AH (H_1): The accident Conditions are different during the 10 week period.

(iii) L.O.S : $\alpha = 0.05$ (5%) assumed

(iv) Test Statistic :

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

$$= 26.6$$

The Critical value of χ^2 at 5% L.of.S with $v = 9$ degree of freedom is 16.9.

(v) Conclusion:

As $\chi^2 > 16.9$, the Null Hypothesis H_0 is rejected.

∴ The accident Conditions are different during the 10 week period.

[part C : 5 Mark]

11. To Compare two kinds of bumper guards, 6 of each kind were mounted on a Car and then the Car was run into a Concrete wall. The following are the Costs of repairs. Use the 0.01 level of Significance to test whether the difference between two Samples mean is Significant?

Guard 1	107	148	123	165	102	119
Guard 2	134	115	112	151	133	129

Sol: Let μ_1 and μ_2 be the population means of Guard 1 and Guard 2.

n_1 = Sample Size of Guard 1 = 6

n_2 = Sample Size of Guard 2 = 6

$\bar{x}_1$ = Sample mean of guard 1

$$= \frac{107+148+123+165+102+119}{6} = \frac{764}{6} = 127.33$$

$\bar{y}_1$ = Sample mean of guard 2

$$= \frac{134+115+112+151+133+129}{6} = \frac{744}{6} = 128$$

x	$(x-\bar{x})$	$(x-\bar{x})^2$	y	$(y-\bar{y})$	$(y-\bar{y})^2$
107	-20.33	413.31	134	5	25
148	20.67	427.25	115	-14	196
128	-4.33	18.75	112	-17	289
165	37.67	1419.03	151	22	484
102	-25.33	641.61	133	4	16
119	-8.33	69.39	129	0	0
		2989.34	774		1010

$$s^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2\right]$$

$$= \frac{1}{6+6-2}\left[2989.34+1010\right] = \frac{1}{10}(3999.34)$$

$$= 399.934$$

$$S = \sqrt{399.934} = 19.998$$

S = Sample S.D = 19.998

ν = degree of freedom = n_1+n_2-2

$= 6+6-2 = 10$

(i) NH (H_0) : $\mu_1 = \mu_2$ i.e, there is no significant difference between two sample means

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

(iii) Level of significance : $\alpha = 0.01$ (1%) ~~assumed~~ given

(iv) Test statistic:

$$t = \frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{127.33-129}{19.998\sqrt{\frac{1}{6}+\frac{1}{6}}}$$

$$= \frac{-1.67}{19.998 \times 0.577} = -0.144$$

The critical value of t at 0.01 L.o.f.s with 10 degree of freedom for two-tailed test is

$$t_{\alpha/2} = t_{0.005} = 3.169$$

(v) Conclusion:

As $|t| < 3.169$, the Null Hypothesis H_0 is accepted

$\therefore$ There is no significant difference between two means.

12. A sample of 26 bulbs gives a mean life of 990 hrs with a S.D of 20 hrs. The manufacturer claims that the mean life of bulbs is 1000 hrs. Is the sample not up to the standard.

Sol: Given,

n = Sample size = 26

$\bar{x}$ = Sample mean = 990 hrs

μ = population mean = 1000 hrs.

S = S.D of Sample = 20 hrs

Degree of freedom = $\nu = n-1 = 26-1$

$= 25$

(i) NH $H_0 : \mu = 1000$ i.e the Sample is upto Standard

(ii) AH $H_1 : \mu < 1000$

(iii) Level of Significance : $\alpha = 0.05$ (5%) assumed.

(iv) Test Statistic:

$$t = \frac{\bar{x} - \mu}{S\sqrt{n-1}} = \frac{990 - 1000}{20/\sqrt{25}}$$

$$= \frac{-\cancel{10} \times 5}{\cancel{20}\ 2} = -2.5 \Rightarrow |t| = 2.5$$

The Critical value of t_α at 5% level of Significance for degree of freedom 25 is for one-tailed test is 1.708.

(v) Conclusion:

$|t| > 1.708$, the Null Hypothesis H_0 is rejected.

$\therefore$ the Sample is not upto Standard.

13. A random Sample of Six Steel beams has a mean Compressive Strength of 58,392 p.s.i with a standard deviation of 648 p.s.i. use this information and the level of Significance 0.05 to test whether the true average Compressive Strength of the Steel from which this Sample Came is 58000 p.s.i. Assume normally?

Sol:- Given, n = Sample Size = 6

$\bar{x}$ = Sample mean = 58,392

μ = population mean = 58,000

S = S.D of Sample = 648

Degree of freedom = $v = n-1 = 6-1 = 5$

(i) NH $H_0 : \mu = 58,000$ i.e, the true average Compressive Strength of the steel from which the Sample Came is 58,000.

(ii) AH $H_1: \mu \neq 58,000$

(iii) Level of Significance : $\alpha = 0.05$ (5%) assumed

(iv) Test statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{58,395 - 58,000}{648/\sqrt{5}}$$

$$= \frac{395 \times \sqrt{5}}{648} = 1.353$$

The Critical Value of $t_{\alpha/2}$ at 5% Level of Significance with degree of freedom 5 is 3.365

(v) Conclusion :

As $|t| < 3.365$, the Null Hypothesis H_0 is accepted.

∴ The true average Compressive Strength of the steel from which the Sample Came is 58,000.

14. A die is thrown 264 times with the following result. Show that the die is biassed. (Given $\chi^2_{0.05} = 11.07$ for 5 d.f).

No. on the die	1	2	3	4	5	6
frequency	40	32	28	58	54	52

Sol: Let, $O_1 = 40$, $O_2 = 32$, $O_3 = 28$, $O_4 = 58$, $O_5 = 54$, $O_6 = 52$ be the observed frequencies.

Then, the expected frequencies of each of the Numbers are.

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

$$E_i = \frac{\sum_{i=1}^{n} O_i}{n} = \frac{264}{6} = 44 \quad \forall_i = 1, 2, 3, 4, 5, 6.$$

∴ $E_1 = 44$, $E_2 = 44$, $E_3 = 44$, $E_4 = 44$, $E_5 = 44$, $E_6 = 44$

O_i	E_i	$(O_i - E_i)$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
40	44	-4	16	0.3636
32	44	-12	144	3.2727
28	44	-16	256	5.8181
58	44	4	196	4.4545
54	44	10	100	2.2727
52	44	8	64	1.4545
				17.6362
264	264			

(i) NH H_0 : The die is unbiased

(ii) AH H_1 : The die is biased

(iii) L.O.S : $\alpha = 0.05$ (5%) assumed

(iv) Test statistic:

$$\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i}$$

$$= 17.6362$$

The critical value of χ^2 at 5% L.O.S and $\nu = 5$ degree of freedom is 11.070

(v) Conclusion:

As $\chi^2 > 11.07$ the Null Hypothesis H_0 is rejected.

$\therefore$ The die is biased.

15. pumpkins were grown under two experimental conditions. Two random samples of 11 and 9 pumpkins. show the samples standard deviations of their weights as 0.8 and 0.5 respectively. Assuming that the weight distribution are normal, test the hypothesis that the true variances are equal.

Sol: Let σ_1^2 & σ_2^2 be the variance of two populations

n_1 = 1st Sample Size = 11

n_2 = 2nd Sample Size = 9

s_1 = 1st Sample S.D = 0.8

s_2 = 2nd Sample S.D = 0.5

$$S_1^2 = \frac{n_1 s_1^2}{n_1-1} = \frac{11(0.8)^2}{11-1} = 0.704$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2-1} = \frac{9(0.5)^2}{9-1} = 0.281$$

The degree of freedom $v_1 = n_1 - 1 = 11 - 1 = 10$

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

(i) NH H_0: $\sigma_1^2 = \sigma_2^2$ i.e the variance of two populations is equal.

(ii) AH H_1: $\sigma_1^2 \neq \sigma_2^2$

(iii) L.O.S : $\alpha = 0.05$ (5%) assumed

(iv) Test Statistic:

$$t = \frac{S_1^2}{S_2^2} \quad (S_1^2 > S_2^2)$$

$$= \frac{0.704}{0.281} = 2.5$$

The Critical Value F at 5% L.O.S with $v_1 = 10$ & $v_2 = 8$ degree of freedom is 3.35

(v) Conclusion:

As $F < 3.35$, the Null Hypothesis H_0 is accepted.

∴ The variance of two populations is equal.

[part D : 10 Marks]

16. a) A random sample of 10 boys had the following I.Q's:

70, 120, 110, 101, 88, 83, 95, 98, 107, 100.

(i) Do these data support the assumption of a population mean I.Q of 100?

(ii) Find a reasonable range in which most of the mean I.Q values of samples of 10 boys lie.

Sol: Given,

n = Sample size = 10

$\bar{x}$ = Sample mean

$$= \frac{70+120+110+101+88+83+95+98+107+100}{10}$$

$= 97.2$

μ = population mean = 100

x	$x-\bar{x}$	$(x-\bar{x})^2$
70	-27.2	5739.84
120	22.8	579.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
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
972		1833.60

We know that,

$$s^2 = \frac{1}{(n-1)} \sum_i (x_i - \bar{x})^2 = \frac{1833.60}{9}$$

$$= 203.73$$

$\therefore$ Standard deviation,

$$S = \sqrt{203.73} = 14.27$$

ν = degree of freedom $= n-1 = 10-1 = 9$

S = Sample S.D $= 14.27$

(a) (i) NH $H_0 : \mu = 100$, i.e the data Support the assumption of a population mean IQ of 100

(ii) AH $H_1 : \mu \neq 100$

(iii) Level of Significance: $\alpha = 0.05$ (5%) assumed.

(iv) Test statistic:

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} = \frac{97.2 - 100}{14.27/\sqrt{9}}$$

$$= \frac{-2.8 \times 3}{14.7} = -0.62$$

$$\therefore |t| = 0.62$$

The Critical value of t at 5% level of Significance with 9 degree of freedom for 2-tail test is

$$t_{\alpha/2} = t_{0.025} = 2.26$$

(v) Conclusion:

As $|t| < 2.26$, the NH H_0 is accepted.

$\therefore$ The data Support the assumption of a population mean IQ of 100.

(b) the 95% Confidence interval is given by

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

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

$$= (87.\ 107.4)$$

$\therefore$ the mean IQ values of Samples of 10 boys will lie in 95% Confidence interval (87. 107.4).

17. The time taken by workers in performing a job by method I and method II is given below. Do these data show that the variances of time distribution from population from which these Samples are drawn differ Significantly?

Method I	20	60	26	27	23	22	–
Method II	27	33	42	35	32	34	38

Sol: Let, σ_1^2 & σ_2^2 be the Variances of two methods

n_1 = Sample Size of methods = 6

n_2 = Sample Size of methods = 7

$\bar{x}$ = Sample Size of method 1

$$= \frac{20 + 60 + 26 + 27 + 23 + 22}{6} = 22.23$$

$\bar{y}$ = Sample mean of method 2

$$= \frac{27 + 33 + 42 + 35 + 32 + 34 + 38}{7} = 34.4$$

Degree of freedom = $v_1 = n_1 - 1 = 6 - 1 = 5$

Degree of freedom = $v_2 = n_2 - 1 = 7 - 1 = 6$

x	$(x-\bar{x})$	$(x-\bar{x})^2$	y	$(y-\bar{y})$	$(y-\bar{y})^2$
20	-2.3	5.24	27	-7.4	54.76
16	-6.3	39.69	33	-1.4	1.96
26	3.7	13.69	42	7.6	57.76
27	4.7	22.09	45	0.6	0.36
23	0.7	0.49	32	-2.4	5.76
22	0.3	0.09	34	-0.4	0.16
			38	3.6	12.96
134		81.34	241		133.72

$$S_1^2 = \frac{\Sigma(x_i-\bar{x})^2}{n_1-1} = \frac{81.34}{5} = 16.26$$

$$S_2^2 = \frac{\Sigma(y_i-\bar{y})^2}{n_2-1} = \frac{133.72}{6} = 22.29$$

(i) NH H_0 : $\sigma_1^2 = \sigma_2^2$ i.e their significance differ between Variances of two populations.

(ii) AH H_1 : $\sigma_1^2 \neq \sigma_2^2$

(iii) L.o.S : $\alpha = 0.05$ (5%) assumed

(iv) Test statistic:

$$t = \frac{S_2^2}{S_1^2} \quad (S_2 > S_1)$$

$$= \frac{22.9}{16.26} = 1.3708$$

The critical value of t at 5% L of S with d.f $v_1 = 5$, $v_2 = 6$ is 4.39.

(v) Conclusion: As the $t < 4.39$, the NH H_0 is accepted.

$\therefore$ There is no ~~se~~ significance between variance of two method.

18. A pair of dice are thrown 360 times and the frequency(Y) of each Sum (x) is indicated below.

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

Would you say that dice are fair on the basis of the chi-square test at 0.05 level of significance?

Sol'n Let, $n =$ Sample size $= 11$

Let, $O_1 = 8$ $\quad O_5 = 44$ $\quad O_9 = 26$

$O_2 = 24$ $\quad O_6 = 65$ $\quad O_{10} = 14$

$O_3 = 35$ $\quad O_7 = 51$ $\quad O_{11} = 14$

$O_4 = 37$ $\quad O_8 = 42$

be the observed frequencies

The probability of getting a sum 2, 3, 4 --- 12 when a pair of dies are thrown is

x_i	2	3	4	5	6	7	8	9	10	11	12
$P(x_i)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

O_i	$E_i = 360 \times p(x_i)$	$(O_i - E_i)$	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
8	10	-2	4	0.4
24	20	4	16	0.8
35	30	5	25	0.833
37	40	-3	9	0.225
44	50	-6	36	0.72
65	60	5	25	0.417
51	50	1	1	0.02
42	40	2	4	0.1
26	30	-4	16	0.53
14	20	-6	36	1.8
14	10	4	16	1.6
360				7.445

(i) NH H_0: The dies are fair

(ii) AH H_1: The dies are not fair

(iii) L.O.S: $\alpha = 0.05$ (5%) given

(iv) Test statistic: $x^2 = \dfrac{\Sigma (O_i - E_i)^2}{E_i} = 7.445$

The critical of x^2 at 5% L.O.S with $V = 10$ degree of freedom is 18.3

(v) Conclusion:

As $x^2 < 18.3$, the Null Hypothesis H_0 is accepted.

∴ The dies thrown are fair.