Unit: 4 - Small Sample tests

[Part : A] [1 marks]

1. (a) Write down the Applications of the F-distribution? [L1 211.4]

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.

(b) what is that Degree of freedom? [L2 211.4]

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 $\nu = n-k$.

(c) Write the applications of Chi-square distribution. [L2 211.4]

* To test the goodness of fit.
* To test the indipendence of attributes.
* To test the homogeneity of independent estimation of the population variance.

(d) write the uses of the t-test. [L1 211.4]

* 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.

(e) Write the formula for Chi-square ~~Distribution~~ test. [L2 211.4]

The chi-square formula is,

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

where O_i = Observed value (actual value) &

E_i = expected value.

[Part : B] [3 marks]

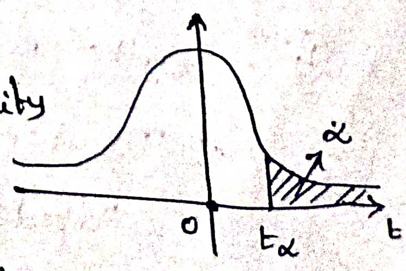
(2) (a) Write down the properties of t-Distribution. [L1 211.4]

(1) The shape of t-distribution is bell-shaped, which is similar to that of a normal distribution and is symmetrical about the mean.

(2) The t-distribution curve is also asymptotic to the t-axis.

(3) It is symmetrical about the line t=0.

(4) The form of the probability curve varies with degrees of freedom.



(5) It is unimodel with Mean = Median = Mode.

(b) A random sample of size 25 from a normal population has the mean $\bar{x} = 47.5$ and the S.D $s = 8.4$. Does this information tend to support or refute the claim that the mean of the population is $\mu = 42.5$? [L3 211.4]

Solu

Gn $n = 25$, $\bar{x} = 47.5$, $\mu = 42.5$, $s = 8.4$.

We have t - distribution is, $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

$\nu = n - 1$
$= 24$

$\alpha = 0.05$

$$t = \frac{47.5 - 42.5}{8.4/\sqrt{25}} \Rightarrow 2.98$$

$\boxed{t = 2.98}$ This value of t has 24 degrees of freedom.

$t_{\alpha}24 = 2.797$. We conclude that the information of given data, the mean of the population is $\mu = 42.5$.

(c) 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 S.D obtained were 17.85 and 1.955 respectively. Is the result of experiment significant? [L4 211.4]

Solu: Gn $n = 14$, $\bar{x} = 17.85$, $s = 1.955$, $\mu = 18.5$, dof $= n - 1$
$= 13$

(i) H_0: The result of the experiment is not significant.
i.e, $\mu = 18.5$.

(ii) H_1: $\mu \neq 18.5$

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

(4) The test statistic is,

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n-1}} = \frac{17.85-18.5}{1.955/\sqrt{13}}$$

$$\boxed{t = -1.199}$$

$\therefore$ $|t| = 1.199$ ie, calculated $t = 1.199$ //.

$t_{\alpha/2}^{13} = 2.16$ (table value)

$\therefore$ $1.199 < 2.16$ we accept the Null hypothesis H_0 at 5% and conculde that the result of the experiment is not ~~ex~~ significant.

(d) 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 are 26.94 and 18.73 respectively. Can the sample be considered to have been drawn from the same normal population. [L3 211.4]

Solu:

Gn. $n_1 = 9$, $n_2 = 7$, $\bar{x} = 196.42$, $\bar{y} = 198.82$ 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) H_0 : The two Samples are drawn from the sample population.

ie, $\mu_1 = \mu_2$.

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

(3) $\alpha = 0.05$

(4) The test statistic is,

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{196.42 - 198.82}{1.81\sqrt{\frac{1}{9} + \frac{1}{7}}}$$

$$t = -2.63$$

Cal. value $|t| = 2.63$.

$t_{\alpha/2}^{14} = 2.15 < 2.63$.

We reject the null hypothesis and conclude that the two samples are not drawn from the same population.

(e.) 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. [LQ 211.4]

Solu: Expected frequence of accidents each week $= \frac{100}{10} \Rightarrow \underline{\underline{10}}$.

(i) H_0 : The accident cond's were the same during the 10 week period.

(ii) H_1 : The " are different during " " " "

~~obse~~ O_i	E_i	$[O - E_i]$	$(O_i - E_i)^2 / E_i$
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

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

$$\boxed{\chi^2 = 26.6}$$

Calculate $\chi^2 = 26.6$

Here $n = 10$ observations are given.

$\therefore$ $d.f = n-1 \Rightarrow 10-9 \Rightarrow 9$

tabulated $\chi^2 = 16.9$

Since Calculated $\chi^2 >$ tabulated χ^2.

$\therefore$ The Null hypothesis is rejected and conclude that the accident conditions were not the same during the 10 week period.

③(a). [5 marks].

③ (a). If two independent random samples of size $n_1 = 13$ and $n_2 = 7$ are taken from a normal population, what is the probability that the variance of the first sample will be atleast four times as large as that of the 2nd sample [LS. 211.4]

Solu:- Gn $n_1 = 13$, $n_2 = 7$.

Then $\nu_1 = n_1 - 1 \Rightarrow 12$ & $\nu_2 = n_2 - 1 \Rightarrow 6$ and

$s_1^2 = 4 s_2^2$

i.e., the variance of the first sample will be at least four times as large as that of the 2nd sample.

Now, $F = \dfrac{s_1^2}{s_2^2} \Rightarrow \dfrac{4 s_2^2}{s_2^2} \Rightarrow 4.00$

This value of F follows F-distribution with $\nu_1 = 12$ and $\nu_2 = 6$ degrees of freedom.

Hence from the table we get $F_{0.05, 12, 6} = 4.00$

$\therefore$ The required probability is 0.05.

(b). 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 upto the standard.

[L3 211.4]

Solu:

Here sample size $n = 26$, $\bar{x} = 990$

$\mu = 1000$, $S. = 20$, $dof = n-1 \Rightarrow 25$.

Here we know $\bar{x}$, μ, S.D & n.

$\therefore$ we use students 't' test.

(i) H_0: The sample is upto the standard.

(2) H_1: $\mu < 1000$.

(3) $\alpha = 0.05$

(4) The test statistic is $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n-1}} \Rightarrow \dfrac{990 - 1000}{20/\sqrt{25}}$

$t = -2.5$

$\therefore |t| = 2.5$ ie., Calculated value of $\boxed{t = 2.5}$

$t_{\alpha}{}^{25} = 1.708$

$\therefore 2.5 > 1.708$. we reject the null hypothesis H_0.

and conclude that the sample is not upto the standard.

(c) The heights of 10 males of a given locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66 inches. Is it reasonable to believe that the average height is greater than 64 inches? Test at 5% significance level assuming that for 9 degrees of freedom ($t = 1.833$ at $\alpha = 0.05$).

[L2 211.4]

Solu

Mean, $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{660}{10}$

$\boxed{\bar{x} = 66}$ $dof = n-1 \Rightarrow 9$

x	$x-\bar{x}$	$(x-\bar{x})^2$
70	4	16
67	1	1
62	-4	16
68	2	4
61	-5	25
68	2	4
70	4	16
64	-2	4
64	-2	4
66	0	0
660		90

w.k.t $s^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$

$= \frac{1}{9} \times 90$

$s^2 = 10$

$\boxed{s = 3.16}$

(i) H_0 : The avg height is not greater than 64 inches.

ie, $\mu = 64$ inches.

(ii) $H_1 : \mu > 64$ inches.

(iii) $\alpha = 0.05$.

(iv) Test statistic is,

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{66-64}{3.16/\sqrt{9}} \Rightarrow 1.9 \quad \therefore \boxed{t = 1.9}$$

$t^{9}_{0.05} = 1.833$ (given)

Cal.val > tab.val. we reject the H_0. we Conclude that the avg. height is greater than 64 inches.

(d). A die is thrown 264 times with the following results. Show that the die is biased [Gn: $\chi^2_{0.05} = 11.07$ for 5 d.f]

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

[L4 · 2.11.4]

Solu:-

H_0 : The die is unbiased.

The expected frequency of each of the numbers 1, 2, 3, 4, 5, 6 is,

$$\frac{264}{6} = \underline{\underline{44}}$$

Calculations for χ^2.

O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
40	44	16	0.3636
32	44	144	3.2727
28	44	256	5.8181
58	44	196	4.4545
54	44	100	2.2727
52	44	64	1.4545
264	264		17.6362

$$\therefore \chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \Rightarrow 17.6362 \text{ //}$$

$$dof = n-1 \Rightarrow 5 \text{ //}$$

$$t_{\alpha}^{5} = 11.07$$

$\therefore$ $17.63 > 11.07$.

we reject the null hypothesis.

i.e., we reject the hypothesis that the die is unbiased.

(e) A stimulus administered to 12 patients resulted in the following changes in the blood pressure : 5, 2, 8, −1, 3, 0, −2, 1, 5, 0, 4, 6. Can it be concluded that the stimulus will in general be accompanied by an increase in blood pressure?

[L5 211.4].

Soln:-

Let x denote the increment in BP.

we shall Compute the mean and S.D of BP as follows.

$$\bar{x} = \frac{\Sigma x_i}{12} = \frac{1}{12}[5+2+8-1+3+0-2+1+5+0+4+6]$$

$$= \frac{31}{12} \Rightarrow 2.58$$

$$s^2 = \frac{1}{n^2}\sum_i (x_i - \bar{x})^2 = \frac{1}{144}\Big[(5-2.58)^2 + (2-2.58)^2 + (8-2.58)^2$$

$$+ (-1-2.58)^2 + (3-2.58)^2 + (0-2.58)^2 + (-2-2.58)^2$$

$$+ (1-2.58)^2 + (5-2.58)^2 + (0-2.58)^2 + (4-2.58)^2 + (6-2.58)^2$$

$$s^2 = 9.53$$

$$\boxed{s = 3.08}$$

dof = n−1
= 12−1
= 11

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

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

(iii) $\alpha = 0.05$

(iv) $$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{2.58 - 0}{3.08/\sqrt{12}} \Rightarrow 2.89.$$

Tab. Val $t^{11}_{\frac{0.05}{2}} = 2.2$ $\therefore$ $2.89 > 2.2$

i.e., Cal. val t > tab. val t $\therefore$ H_0 is rejected at 5%.

we may conclude that the stimulus will in general be accompanied by an increase in BP.

(4)(a) A random Sample of 10 boys had the following I.Q's : 70, 120, 110, 101, 88, 83, 95, 98, 107 and 100.

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

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

[L5 211.4]

Solu:-

(a) Here S.D. and mean of the sample is not given directly. we have to determine these S.D. and mean as follows.

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

x	$x-\bar{x}$	$(x-\bar{x})^2$
70	-27.2	739.84
120	22.8	519.84
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

w.k.t $s^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$

$= \frac{1833.60}{9}$

$s^2 = 203.73$

$s = \sqrt{203.73}$

$\boxed{s = 14.27}$

(1) H_0 : The data support the assumption of a population mean I.Q of 100 in the population. i.e., $\mu = 100$.

(2) H_1 : $\mu \neq 100$.

(3) $\alpha = 0.05$

(4) The test statistic is, $t = \frac{\bar{x} - \mu}{S/\sqrt{n}} = \frac{97.2 - 100}{14.27/\sqrt{10}}$

$$\boxed{t = -0.62}$$

$\therefore$ $|t| = 0.62$ ie., Calculated value of $t = 0.62$.

$t^{9}_{\alpha/2} = 2.26$ dof $= n-1$
$= 10-1$
$= 9$

$2.26 > 0.62$.

Cal. Val of t < tab. Valu of t. we accept the null hypothesis H_0.

ie., the data support. the assumption of mean I.Q of 100 in the population.

(b) The 95% Confidence limits are given by,

$$\bar{x} \pm t_{0.05}\, S/\sqrt{n}$$

$$= 97.2 \pm 2.26 \times 4.512$$

$$= 97.2 \pm 10.198$$

$$= 107.4 \text{ \& } 87.$$

$\therefore$ the 95% Confidence limits with in which the mean I.Q values of samples of 10 boys will lie is (87, 107, 40).

(b) 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 populations have the same variance.

[LB 211.4]

Solu :-

Let σ_1^2 & σ_2^2 be the variances of the two normal populations from which the samples are drawn.

(i) $H_0 : \sigma_1^2 = \sigma_2^2$

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

Here $n_1 = 8$, $n_2 = 10$

Also, $\Sigma (x_i - \bar{x})^2 = 84.4$, $\Sigma (y_i - \bar{y})^2 = 102.6$

If S_1^2 & S_2^2 be the estimates of σ_1^2 & σ_2^2 then,

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

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

Let H_0 be true. Since $S_1^2 > S_2^2$,

The test statistic is,

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

Cal.val $\boxed{F = 1.057}$

dof $\nu_1 = n_1 - 1 = 8 - 1 \Rightarrow 7.$

$\nu_2 = n_2 - 1 = 10 - 1 \Rightarrow 9.$

$$\boxed{F_{0.05}(7,9) = 3.29}$$

$\therefore$ $1.057 < 3.29$

Cal.val F < tab.value F, we accept the Ho and conclude that the populations have the same variance.

(C) A pair of dice are thrown 360 times 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 dice are fair on the basis of the chi-square test at 0.05 level of significance?

[L3 211.4]

Solu:

(1) H_0 : The dice are fair.

(2) H_1 : The dice are not fair.

(3) level of significance : $\alpha = 0.05$

The probabilities of getting a sum 2,3,4,5,6,7, 8,9,10,11, and 12 are

$X = x_i$:	2	3	4	5	6	7	8
$P(x_i)$:	1/36	2/36	3/36	4/36	5/36	6/36	5/36

9	10	11	12
4/36	3/36	2/36	1/36.

Calculations for χ^2.

Sum	O_i	$E_i = 360 \cdot P(x)$	$(O-E_i)^2$	$(O_i-E_i)^2/E_i$
2	8	10	4	0.4
3	24	20	16	0.8
4	35	30	25	0.833
5	37	40	9	0.225
6	44	50	36	0.72
7	65	60	25	0.417
8	51	50	1	0.02
9	42	40	4	0.1
10	26	30	16	0.53
11	14	20	36	1.8
12	14	10	16	1.6
	360	360		7.445

$$\therefore \chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i} = 7.445.$$

dof = n-1 => 11

tab.val χ^2 for 11 dof at 5% is 18.3

$\therefore 7.445 < 18.3$

ie, Cal.Value χ^2 < tab.val, we accept the H_0.

$\therefore$ The dice are fair.