CMR INSTITUTE OF TECHNOLOGY

USN

Internal Assessment Test – II 2024

Sub:	**Mathematics for Computer Science**							Code:	BCS301
Date:	13/12/24	Duration:	**90 mins**	Max Marks:	**50**	Sem :	**3**	Branch:	**CSE / ISE / AI-DS / AIML**

Answer Any 5 Questions.

		Marks	OBE CO	OBE RBT
1	Analyze and interpret the following statistics concerning output of wheat per field obtained as a result of experiment conducted to test four varieties of wheat viz., A, B, C and D under a Latin-Square design. **[Given: F(3, 6) = 4.76 at 5% l.o.s]**	[10]	CO6	L3
2	Set up an analysis of variance table for the following two-way design results: per acre production data of wheat in metric tons. Also state whether variety differences are significant at 5% level. **[Given: F(3, 6)=4.76 & F(6, 2)=19.33]**	[10]	CO5	L2
3	In a sample of 600 men from a certain city, 450 are found smokers. In another Sample of 900 men from another city, 450 are smokers. Do the indicate that the Cities are significantly different with respect to the habit of smoking among men. Test at 5% significance level. **[Given Z = 6.38 at 5% l.o.s]**	[10]	CO3	L3
4	a) The nine items of a sample have the following values: 45, 47, 50, 52, 48, 47, 49, 53, 51. Does the mean of these differ significantly from the assumed mean of 47.5 at 5% significance level? **[Given: t= 2.31 for d.f = 8 & 5% l.o.s]** b) Let the observed value of the mean $\overline{X}$ of a random sample of size 30 from a Normal Distribution with mean μ and variance $\sigma^2 = 80$ be 81.2. Find a 90% and a 95% confidence intervals for μ. **[Given Z = 1.96 at 95% C.I & Z= 1.645 at 90%]**	[5+5]	CO3	L3
5	State Central Limit Theorem . An unknown distribution has mean of 90 and a standard deviation of 15. Sample of size n=25 are drawn randomly from the population. Find the probability that the sample mean lies between 85 and 92. **[Given: Z(−1.67)=0.4514, Z(0.67)= 0.2454]**	[10]	CO4	L3
6	A sample analysis of the examination of 500 students was made. It was found that 220 students had failed, 170 had secured third class, 90 had secured second class and 20 had secured first class. Do the figures support the general examination result which is in ratio 4:3:2:1 for the respective categories **(Chi-square test value at 5% l.o.s & 3 d.f is 7.81)**	[10]	CO4	L3

Question 1 data:

C-25	B-23	A-20	D-20
A-19	D-19	C-21	B-18
B-19	A-14	D-17	C-20
D-17	C-20	B-21	A-15

Question 2 data:

	Varieties of seeds		
Varieties of fertilizers	A	B	C
W	6	5	5
X	7	5	4
Y	3	3	3
Z	8	7	4

Maths - IAT 2

Q1.

C25	B23	A20	D20
A19	D19	C21	B18
B19	A14	D17	C20
D17	C20	B21	A15

By coding method, subtract 20 from all values to simplify calculations.

row/column	1	2	3	4	T_i	T_i^2
1	C_5	B_3	A_0	D_0	8	64
2	A_{-1}	D_{-1}	C_1	B_{-2}	-2	4
3	B_{-1}	A_{-6}	D_{-3}	C_0	10	100
4	D_{-3}	C_0	B_1	A_{-5}	-7	49
T_j	0	-4	-1	-7	T = -12	
T_j^2	0	16	1	49		

Correction factor $C.f = \frac{T^2}{N} = \frac{(-12)^2}{16} = \frac{144}{16} = 9$

$TSS = \sum x_{ij}^2 - Cf = 122 - 9 = 113$

$SSR = \frac{T_i^2}{n_i} - Cf = \frac{64}{4} + \frac{4}{4} + \frac{100}{4} + \frac{49}{4} - 9$

$= 16 + 1 + 25 + 12.25 - 9$

$= 46.25$

$SSC = \frac{T_j^2}{n_j} - Cf = \frac{0}{4} + \frac{16}{4} + \frac{1}{4} + \frac{49}{4} - 9$

$= 0 + 4 + 0.25 + 12.25 - 9 = 7.5$

Rearranging the coded data by variety:

Treatments	Obj-1	Obj-2	Obj-3	Obj-4	Tv
A	-1	-6	0	-5	-12
B	-1	3	1	-12	-9
C	5	0	1	1	7
D	-3	-1	6	-3	-1

$SSL = \frac{Tv^2}{n_v} - CF = \frac{144}{4} + \frac{1}{4} + \frac{36}{4} + \frac{49}{4} - 9$

$= 36 + 0.25 + 9 + 12.25 = 48.5$

$SSE = TSS - (SSC + SSR + SSL)$

$= 133 - (7.5 + 46.25 + 48.5) = 113 - 102.25 =$

$= 10.75$

df for rows = $r - 1 = 3$

df for columns = $c - 1 = 3$

df for varieties = $V - 1 = 3$

df for residuals = $(c-1)(r-2) = 6$

df for total = $n - 1 = 15$

ANOVA Table:

Source of variation	SS	df	MSS	F-Ratio
Rows	46.50	3	15.50	8.85
Columns	7.50	3	2.50	1.43
Varieties/Treatments	48.50	3	16.17	9.24
Residual (Error)	10.50	6	1.75	–
Total	113.00	15	–	–

Conclusion:- Variance between rows ($F_R = 8.85$) and varieties ($F_V = 9.24$) are significant ($F > 4.76$).

Variance between columns ($F_C = 1.43$) is not significant ($F < 4.76$). Row effects & variety effects influence yeild but column ~~affects~~ do not.

Q2.

plot	Varieties A	B	C	Row total
1	6	5	5	16
2	7	5	4	16
3	3	3	3	9
4	8	7	4	19
Column total	24	20	16	20

Null hypothesis H_0: The mean production for the three wheat varieties does not differ.

~~H~~ H_1 : The mean production for the four plots does not differ.

$$CF = \frac{T^2}{N} = \frac{60^2}{12} = 300$$

$$TSS = \sum X_{ij}^2 - CF = 6^2+5^2+5^2+7^2+5^2+4^2+3^2+3^2+3^2+8^2+7^2+4^2-300$$

$$= 32$$

$SSR = \sum_i \frac{T_i^2}{n_i} - CF = \frac{16^2}{3} + \frac{16^2}{3} + \frac{9^2}{3} + \frac{19^2}{3} - 300$

$= 318 - 300 = 18$

$SSC = \sum_j \frac{T_j^2}{n_j} - CF = \frac{24^2}{4} + \frac{20^2}{4} + \frac{16^2}{4} = 300$

$= 308 - 300 = 8$

$SSE = TSS - (SSR + SSC) = 32 - (18+8) = 6$

ANOVA Table:-

Source	SS	df	MSS	f-ratio
Rows (Plots)	18	3	18/3 = 6	6/1 = 6
Columns (varieties)	8	2	8/2 = 4	4/1 = 4
Error	6	6	6/6 = 1	
Total		11		

Conclusion:-

For columns:- $F_c = 6.98 > 5.14$. Reject H_0. The wheat varies significantly.

Q3. Null H_0 : Proportion of smokers in the two cities is the same $(P_1 = P_2)$

H_1 : The proportion of smokers in the 2 cities is different $(P_1 \neq P_2)$

Given:- $n_1 = 600 \quad x_1 = 450$

$\hat{p_1} = \frac{x_1}{n_1} = \frac{450}{600} = 0.75$

$n_2 = 900 \quad x_2 = 450$

$\hat{p_2} = \frac{x_2}{n_2} = \frac{450}{900} = 0.5$

$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{450+450}{600+900} = \frac{900}{1500} = 0.6$

$q = 1 - p = 1 - 0.6 = 0.4$

Z-test for difference of proportions:-

$$Z = \sqrt{\frac{p_1 - p_2}{p \cdot q \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \sqrt{\frac{0.75 - 0.5}{0.6 \times 0.4 \times \left(\frac{1}{600} + \frac{1}{900}\right)}} = 9.69$$

$Z = 9.69 > 1.96$ ∴ null hypothesis is rejected.

Q4. a) $H_0 = \mu = 4.75$

$H_1 = \mu \neq 4.75$

$$\bar{x} = \frac{45+47+50+52+48+47+49+53+51}{9}$$

$$= \frac{442}{9} = 49.11$$

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

$s = \sqrt{\frac{54.88}{9-1}} = \sqrt{\frac{54.88}{8}} = \sqrt{6.86} \approx 2.62$

$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{49.11 - 47.5}{2.62/\sqrt{9}} = \frac{1.61}{2.62/3}$

$= \frac{1.61}{0.87} \approx 1.85$

$df = n-1 = 9-1 = 8$

$|t| = 1.85 < 2.31$, we fail to reject the null hypothesis.

Q4 b) Given $\bar{x} = 81.2$

$n = 30 \quad \sigma^2 = 80 \quad \sigma = \sqrt{80} = 8.94$

Confidence Interval $= \bar{x} \pm Z \cdot \frac{\sigma}{\sqrt{n}}$

$\frac{\sigma}{\sqrt{n}} = \frac{8.94}{\sqrt{30}} \approx \frac{8.94}{5.477} \approx 1.63$

For 90% Confidence Interval $(Z = 1.645)$

$CI = 81.2 \pm (1.645 \times 1.63) = (78.52, 83.88)$ //

For 95% Confidence Interval $(Z = 1.96)$

$CI = 81.2 \pm (1.96 \cdot 1.63) = (78.00, 84.40)$ //

Q5. Central Limit Theorem : Let $x_1, x_2, \ldots x_n$ be a random sample of size n drawn from a population with mean μ and variance σ^2. As the sample size n becomes large, the sampling distribution of the sample mean $\bar{x}$ approaches a normal distribution with mean μ & variance σ^2/n, regardless of the shape of the original population distribution.

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \text{ for large } n.$$

Given:- $\mu = 90$, $\sigma = 15$, $n = 25$

$$\bar{x} \sim N\left(\mu = 90, \frac{\sigma^2}{n} = \frac{15}{\sqrt{25}} = 3\right)$$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

for $\bar{x} = 85$:

$$Z_1 = \frac{85 - 90}{3} = -1.67$$

$\bar{x} = 92$:

$$Z_2 = \frac{92 - 90}{3} = 0.67$$

$\therefore$ $P(85 < \bar{x} < 92) = P(-1.67 < Z < 0.67)$

$= P(-1.67 < Z < 0) + P(0 < Z < 0.67)$

$= 0.4514 + 0.2454 = 0.6965$

The probability that the sample mean $\bar{x}$ lies b/w 85 & 92 is

$p(85 < \bar{x} < 92) = 0.6965$ //

Q6. Given: $N = 500$ Ratio: 4:3:2:1

$4+3+2+1 = 10$

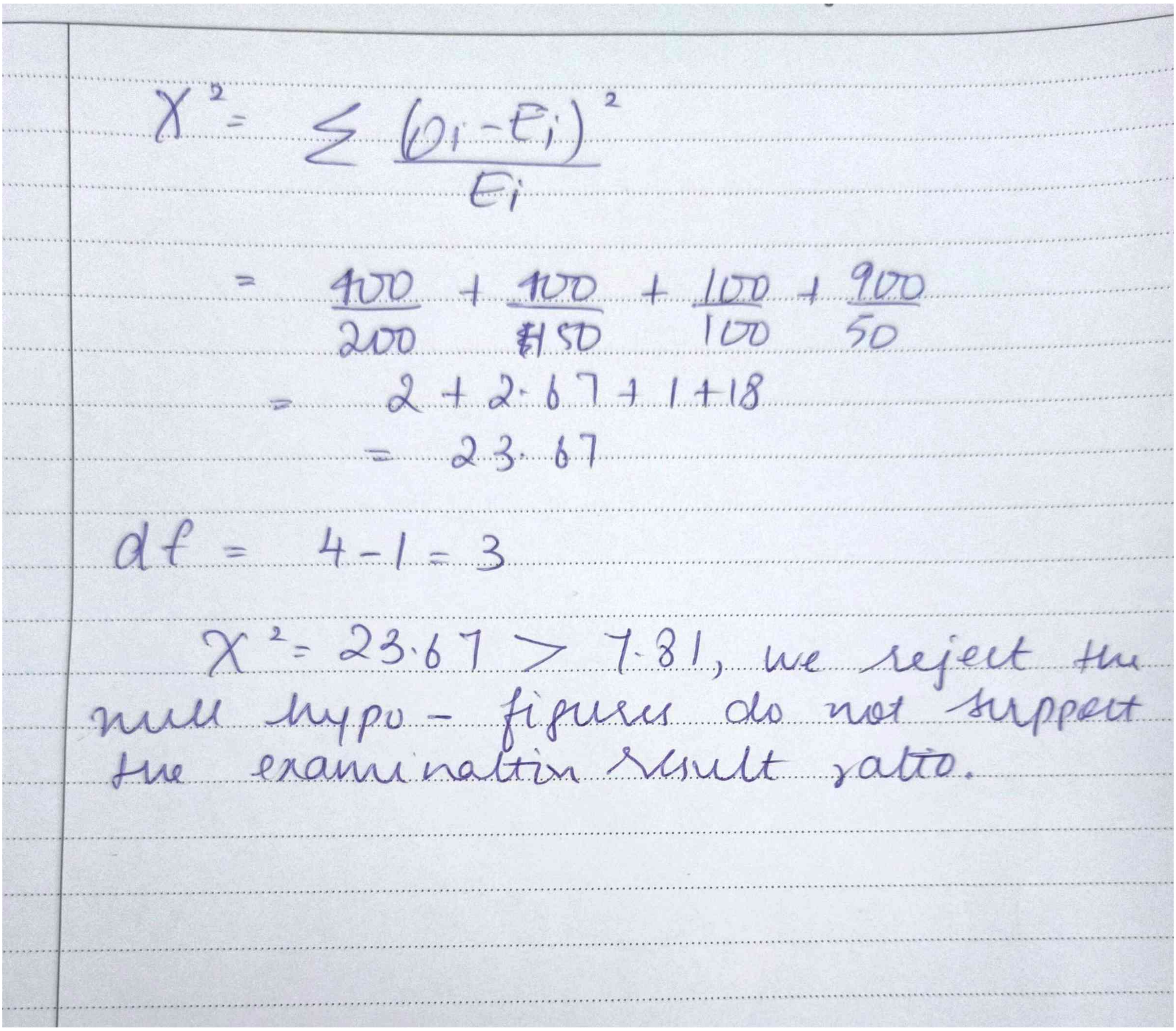
Failed : $\frac{4}{10} \times 500 = 200$

3rd Class $= \frac{3}{10} \times 500 = 150$

2nd Class $= \frac{2}{10} \times 500 = 100$

1st Class $= \frac{1}{10} \times 500 = 50$

Category	O	E	$(O-E)^2$
Failed	220	200	400
3rd Class	170	150	400
2nd Class	90	100	100
1st Class	20	50	900



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

$$= \frac{400}{200} + \frac{400}{150} + \frac{100}{100} + \frac{900}{50}$$

$$= 2 + 2.67 + 1 + 18$$

$$= 23.67$$

$$df = 4 - 1 = 3$$

$\chi^2 = 23.67 > 7.81$, we reject the null hypo - figures do not support the examination result ratio.