Roll No. $\square\square\square\square\square\square\square\square\square\square\square\square$ **Total No. of Pages : 03**

Total No. of Questions : 08

B.Tech (All) (Sem.–2)

MATHEMATICS-II

Subject Code : BTAM204-18

M.Code : 76257

Date of Examination : 15-07-21

Time : 2 Hrs. **Max. Marks : 60**

INSTRUCTIONS TO CANDIDATES :

1. **Attempt any FIVE question(s), each question carries 12 marks.**

1. a) Calculate the coefficient of skewness from the following data :

Life time (in hours)	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No. of tubes	2	5	7	13	21	16	8	3

b) The first four moments for a distribution about the value 4 of the variable are –1.5, 17, –30 and 108. Find the moments about the mean.

2. a) From a bag containing 20 tickets, numbered from 1 to 20, two tickets are drawn at random. Find the probability that

i) Both the tickets have prime numbers on them.

ii) On one there is a prime number and on the other there is multiple of 4.

b) A random variable X has the following probability distribution

X	0	1	2	3	4	5	6	7
P (X)	0	K	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

Determine (i) k (ii) P (X < 3).

3. a) Fit a binomial distribution for the following data :

x	0	1	2	3	4	5
f	2	14	20	34	22	8

b) If 5% of the electric bulbs manufactured by a company are defective.

Use a Poisson distribution to find the probability that in a sample of 100 bulbs.

i) None is defective ii) 5 bulbs will be defective

4. a) Calculate rank correlation coefficient for the following distribution

X	11	6	9	13	6	27	15	16	17	10	1
Y	21	30	37	40	29	34	39	24	20	40	38

b) Find the two regression equations from the following data :

X	1	2	3	4	5
Y	2	3	4	5	6

5. a) A continuous random variable X has the following distribution function.

$$f(x) = \begin{cases} 0, & \text{if } x \le 1 \\ k(x+1)^4, & \text{if } 1 < x \le 3 \\ 1, & \text{if } x > 3 \end{cases}$$

b) Find the value of k and also the probability density function (p.d.f) $f(x)$.

6. Explain the method of least squares to fit a curve. Use this method of find the curve $y = ax + bx^2$ that best fits the following data :

x	1	2	3	4	5
y	1.8	5.1	8.9	14.1	19.8

7. a) A sample of 1000 students from a university was taken and their average weight was found to be 112 pounds with S.D. of 20 pounds. Could the mean weight of students in the population be 120 pounds?

b) The average income of persons was Rs. 210 with S.D. of Rs.10 in a sample of 100 people of a city. For another sample of 150 persons, the average income was Rs. 220 with S.D. of Rs.12. Test whether there is any significant difference between the average incomes of the localities.

8. a) The sales in a supermarket during a week are given below. Test the hypothesis that the sales do not depend on the day of the week, using a significant level of 0.05.

Days	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.
Sales (in Rs.)	65	54	60	56	71	84

b) Discuss the various properties of χ^2-test.

Note: Any student found attempting answer sheet from any other person(s), using incriminating material or involved in any wrong activity reported by evaluator shall be treated under UMC provisions.

Student found sharing the question paper(s)/answer sheet on digital media or with any other person or any organization/institution shall also be treated under UMC.

Any student found making any change/addition/modification in contents of scanned copy of answer sheet and original answer sheet, shall be covered under UMC provisions.

END TERM EXAMINATION

B. Tech CSE / EE / ECE / ME

Mathematics - II

Subject Code: BTAM 204-18

M. Code: 76257

Date of Examination: 15-7-21

Q1 Calculate the Coefficient of skewness from the following

ⓐ data:

Life Time (hours)	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No. of Tubes	2	5	7	13	21	16	8	3

ⓑ The first four moments for a distribution about the value 4 of the variables are -1.5, 17, -30, 108. find the moments about mean.

Solution

ⓐ We can find the coefficient of skewness by using Karl pearson's method

$$S_k = \frac{3(\bar{x} - M)}{S},$$

where $\bar{x}$ = mean of the distribution

M = median of the distribution

S = Standard deviation

Life Time	No of Tubes	Mid Point x_i	fx_i
0-5	2	2.5	5
5-10	5	7.5	35
10-15	7	12.5	87.5
15-20	13	17.5	227.5
20-25	21	22.5	472.5
25-30	16	27.5	440
30-35	8	32.5	260
35-40	3	37.5	112.5
	$\Sigma f = 82$		$\Sigma fx_i = 1670$

Calculating the Mean ($\overline{X}$)

$$\overline{X} = \frac{\Sigma f x_i}{\Sigma f_i} = \frac{1670}{82} = 20.37$$

Now Standard deviation can be calculated as:

$$S = \sqrt{\frac{\Sigma f(x_i - \overline{x})^2}{\Sigma f}} = 8.05$$

To Calculate the Median (M)

we find Median class (where the cf just exceeds half of the total frequency)

So Here Median class is the 4th class (15-20 hrs) with a cummulative frequency of 27

So $$M = L + \frac{\frac{n}{2} - F}{f} \times w$$

$L \rightarrow$ Lower class boundary of median class

$F \rightarrow$ is the cummutative frequency of class before the median class

$f \rightarrow$ frequency of the median class

$w \rightarrow$ width of the median class

$$M = 15 + \frac{\frac{82}{2} - 14}{13} \times 5$$

$$= 15 + \frac{41 - 14}{13} \times 5$$

$$= 15 + 2.08$$

$$\approx 17.08$$

So $$S_k = \frac{3(\overline{x} - M)}{S} \approx \frac{3(20.37 - 17.08)}{8.05} = 0.09$$

So Coeff is tue Hence it indicates that the distribution is truly skewed.

(b) The moment about origin are given

$\mu_1' = -1.5$

$\mu_2' = 17$

$\mu_3' = -30$

$\mu_4' = 108$

Moment about mean $= \mu_1 = \mu_1' - \mu_1' = 0$

(Sample variance) $\mu_2 = \mu_2' - (\mu_1')^2 = (17) - 2.25 = 14.25$

$\mu_3 = \mu_3' - 3\mu_2'(\mu_1') + 2(\mu_1')^3$

Here μ_1, μ_2, μ_3 are 1st, 2nd, 3rd moment about mean

$\mu_3 = (-30) - 3(17)(-1.5) + 2(-1.5)^3$

$= -30 + 76.5 - 6.75$

$= 76.5 - 36.75$

$\mu_3 = 39.25$

Q2 (a) From a bag Containing 20 tickets, numbered from 1 to 20, two tickets are drawn at random. Find the probability that

(i) Both the tickets have prime numbers on them

(ii) on one there is a prime number and on the other there is multiple of 4.

Soln (a)

(i) Both Tickets have prime numbers:

The prime numbers between 1 and 20 are 2, 3, 5, 7, 11, 13, 17 and 19. (8 prime no's) we know $C_{n_r} = \frac{n!}{r!(n-r)!}$

So to choose 2 prime numbered tickets out of 8:

$$^8C_2 = \frac{8!}{2!(8-2)!} = \frac{8!}{2!\,6!} = \frac{8\cdot 7}{2\cdot 1} = 28$$

To choose 2 tickets out of 20:

$$^{20}C_2 = \frac{20!}{2!(20-2)!} = \frac{20!}{2!\,18!} = \frac{20\cdot 19}{2\cdot 1} = 190$$

So $P(\text{Both prime}) = \frac{28}{190}$

(ii) <u>one prime number and one multiple of 4</u>:

no. of prime numbers b/w 1 and 20 = 8

no. of multiples of 4 b/w 1 and 20 = (4, 8, 12, 16) 4

no. of ways to choose 1 prime numbered ticket and 1 multiple of 4:

$$^8C_1 \times {}^4C_1 = 8 \times 4 = 32$$

So, Total number of ways to choose 2 tickets out of 20

$$^{20}C_2 = 190$$

So $P\left(\begin{matrix}\text{one prime and}\\ \text{one multiple}\\ \text{of } 4\end{matrix}\right) = \frac{32}{190}$

(b) A random variable X has the following probability distribution:

X	0	1	2	3	4	5	6	7
P(X)	0	K	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

Determine

(i) K

(ii) $P(X<3)$.

Soln: (i) we know that sum of all probabilities in a probability distribution should be equal to 1

so $\quad \sum P(x) = 1$

$$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

$$12k^2 + 9k = 1$$

$$12k^2 + 9k - 1 = 0$$

$$k = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$k = \frac{-9 \pm \sqrt{9^2 - 4(12)(-1)}}{2(12)}$$

$$k = \frac{-9 \pm \sqrt{117}}{24} = \frac{-9 \pm 10.8}{24}$$

so Either $k = \frac{-9 + 10.8}{24}$ or $\frac{-9 - 10.8}{24}$

$k = 0.075$ or -0.825

Since, $0 < k < 1$ so $\underline{k = 0.075}$

(ii) $\underline{P(x < 3)}$

$$P(x<3) = P(x=0) + P(x=1) + P(x=2)$$

$$P(x<3) = 0 + k + 2k + 2k$$

$$= 5k = 5(0.075)$$

$$P(x<3) = \underline{0.375}$$

Q3 (a) fit a binomial distribution for the following data:

x	0	1	2	3	4	5
f	2	14	20	34	22	8

Soln

To fit a binomial distribution to given data we determine the parameters of binomial distribution

$n \to$ no of Trials

$p \to$ probability of success on a single Trial

The Probability mass function (PMF) of a binomial distribution is given by:

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

where 'k' is the number of successes.

x represents number of successes (0 to 5)

f represents frequency corresponding to 'x'

Total no of trials = $N = \Sigma f$

$$N = 2+14+20+34+22+8 = 100$$

we can fit the binomial distribution by estimating 'p' and 'n' based on above data. The Expected value (μ) of a binomial distribution is 'np'.

Variance (σ^2) is $np(1-p)$

$$\mu = np$$

$$\sigma^2 = np(1-p)$$

$$\hat{p} = \frac{\Sigma xf}{N} = \frac{230}{100} = 2.3$$

$$\hat{n} = \frac{N}{\hat{p}} = \frac{100}{2.3} = 43.48$$

So Binomial distribution would be $B(43, 0.23)$

Q3(b) If 5% of Electric bulbs manufactured by a Company are defective. use Poisson dist to find the probability that in sample of 100 bulbs.

(i) None is defective

(ii) 5 bulbs will be defective.

Soln

For Poisson distribution, the probability mass function is given by:

$$P(X=k) = \frac{e^{-\lambda}\lambda^{k}}{k!}$$

where X is the number of events, 'k' is the number of occurences and λ is the average rate of occurence

Average rate of defective bulbs (λ) = Product of probability of single bulb (p) × Total number of bulbs in the sample (n)

$$\lambda = np$$

Here $p = 0.05$, $n = 100$

$$\lambda = 100 \times 0.05 = 5$$

(i) None is defective: (k=0)

$$P(X=0) = \frac{e^{-5} \times 5^{0}}{0!} = e^{-5}$$

(ii) 5 bulbs will be defective (k=5)

$$P(X=5) = \frac{e^{-5} \times 5^{5}}{5!}$$

Q4 (a) Find rank correlation coefficient for the following distribution

X	11	6	9	13	6	27	15	16	17	10	1
Y	21	30	37	40	29	34	39	24	20	40	38

Soln

First we rank the data for X and Y

X : 7, 2, 4, 9, 2, 11, 6, 8, 10, 5, 1

Y : 6, 8, 10, 11, 7, 9, 10, 5, 3, 11, 10

Now, difference in rank pair (d)

d : -2, 0, -1, -2, 5, 2, 0, 3, 1, -6, -9

d^2 : 4, 0, 1, 4, 25, 4, 0, 9, 1, 36, 81

Sum d^2 (Σd^2) = 165

$n = 11$ (no of data points)

Rank Correlation Coefficient:

$$r_s = 1 - \frac{6 \Sigma d^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6 \times 165}{11(11^2-1)} = 1 - \frac{990}{1100} = 1 - 0.9 = 0.1$$

So rank correlation coefficient is 0.1

(b) Find Two regression equations from the following data:

X	1	2	3	4	5
Y	2	3	4	5	6

Soln: To find the regression equations we will calculate the slope (b) and the intercept (a) for both the regression lines (Y on X & X on Y) using least square method.

Regression Equation of Y on X:

(1) Calculate the means $\bar{X} = \frac{\Sigma X}{n} = \frac{1+2+3+4+5}{5} = 3$

$\bar{Y} = \frac{\Sigma Y}{n} = \frac{2+3+4+5+6}{5} = 4$

The value of slope (b) is given by

$$b = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

$$= \frac{(1-3)(2-4) + (2-3)(3-4) + (3-3)(4-4) + (4-3)(5-4) + (5-3)(6-4)}{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}$$

$$b = \frac{-2-1+0+1+4}{4+1+0+1+4} = \frac{2}{5}$$

and Intercept 'a': $a = \bar{Y} - b\bar{X} = 4 - \frac{2}{5} \times 3$

$$= \frac{12}{5}$$

So the regression equation for Y on X is $\boxed{Y = \frac{2}{3}X + \frac{12}{5}}$

Regression Equation of X on Y:

calculating the the slope 'b'

$$b = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

$$b = \frac{(2-4)(1-3) + (3-4)(2-3) + (4-4)(3-3) + (5-4)(4-3) + (6-4)(5-3)}{(2-4)^2 + (3-4)^2 + (4-4)^2 + (5-4)^2 + (6-4)^2}$$

$$b = \frac{2+1+0+1+4}{4+1+0+1+4} = \frac{8}{10} = \frac{4}{5}$$

model or curve.

and Intercept 'a'

$$a = \bar{X} - b\bar{Y} = 3 - \frac{4}{5} \times 4 = 8/5$$

So the regression Equation for X on Y is $\boxed{X = \frac{4}{5}Y + \frac{8}{5}}$

Q5 (a) A Continuous random variable X has the following distribution function

$$f(x) = \begin{cases} 0 & , \text{ if } x \le 1 \\ k(x+1)^4 & , \text{ if } 1 < x \le 3 \\ 1 & , \text{ if } x > 3 \end{cases}$$

(b) Find the value of 'k' and also find the probability density function (p.d.f)(f(x)).

Soln:

We know that total area under the probability density function must be equal to 1. Therefore we have

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow \int_{-\infty}^{1} 0\,dx + \int_{1}^{3} k(x+1)^4 dx + \int_{3}^{\infty} 1 \cdot dx = 1$$

$$\Rightarrow 0 + \int_{1}^{3} k(x+1)^4 dx + 0 = 1$$

$$\Rightarrow \int_{1}^{3} k(x+1)^4 dx = 1$$

$$\Rightarrow k \left| \frac{(x+1)^5}{5} \right|_{1}^{3} = 1$$

$$\Rightarrow k \left[\frac{(3+1)^5}{5} - \left(\frac{(2)^5}{5} \right) \right] = 1$$

$$\Rightarrow k \left[\frac{1024}{5} - \frac{32}{5} \right] = 1$$

$$\Rightarrow k \left[\frac{992}{5} \right] = 1$$

$$\Rightarrow k = \left[\frac{5}{992} \right] = 0.005$$

To find the p.d.f. $f(x)$:

Now we differentiate CDF to get p.d.f.

$$f(x) = \frac{d}{dx} F(x)$$

where $F(x)$ is Cummulative distribution fn (CDF)

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$f(x) = \frac{d}{dx}\left[\int_{-\infty}^{x} f(t)\,dt\right]$$

$$f(x) = \frac{d}{dx}\left[5k\int_{-\infty}^{x} (\cancel{t}+1)^4 dt\right)$$

$$= 5k \frac{d}{dt}\int_{-\infty}^{x}\left[(t+1)^4 dt\right]$$

$$= 5k \frac{d}{dt}\left(\frac{t+1)^5}{5}\right|_{-\infty}^{x} dt$$

$$= \cancel{5}k\frac{d}{dt}\left[\frac{(x+1)^5}{\cancel{5}}\right]$$

$$= k \cdot 5(x+1)^4$$

$$= 0.005 \times 5\,(x+1)^4$$

$$= 0.025\,(x+1)^4.$$

Q6 Explain the method of least squares to fit a curve. use this method to find the curve $y = ax + bx^2$ that best fits the following data:

x	1	2	3	4	5
y	1.8	5.1	8.9	14.1	19.8

Soln:

The method of least squares is a statistical approach used to find the best-fitting curve to a set of data points. we aim to minimize the sum of the squared differences between the observed values and the values predicted by the model or curve.

For the given data points (x_i, y_i) where $i = 1,2,3,\ldots n$ and a model function $y = f(x,a,b)$, the sum of squared differences is:

$$S(a,b) = \sum_{i=1}^{n} [y_i - f(x_i, a, b)]^2$$

we find values of 'a' and 'b' which minimize $S(a,b)$.

Here $y = ax + bx^2$

so $\frac{\partial S}{\partial a} = 0 \Rightarrow -2 \sum_{i=1}^{n} x_i(y_i - (ax_i + bx_i^2))$

$\frac{\partial S}{\partial b} = 0 \Rightarrow -2 \sum_{i=1}^{n} x_i^2 (y_i - (ax_i + bx_i^2))$

On solving we get

$a = 1.07$ and $b = 0.85$

So the best fitting curve is, $y = 1.07x + 0.85x^2$

Q7

(a) A sample of 1000 students from a university was taken and their average weight was found to be 112 pounds with SD of 20 pounds. Could the mean weight of students in the population be 120 pounds?

Soln

We use Hypothesis Testing

The null Hypothesis (H_0) : There is no significant difference in mean weight (120 pounds)

Alternative hypothesis (H_1) : There is a significant difference, population weight is 120 pounds.

Here Sample mean $\bar{x} = 112$ pounds

(s) S.D = 20 pounds

(n) Sample size = 1000

We use Z- Test

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

$$Z = \frac{112 - 120}{\frac{20}{\sqrt{1000}}} = \frac{-8}{\frac{20}{\sqrt{1000}}} = -2.83$$

For Two tailed test at 0.05 significance level the critical value of Z are -1.96 and 1.96

So, since the calculated score is less than Tabular value (-1.96), we Reject the Null hypothesis

Hence, at 0.05 level of significance the mean weight of students in the population is not 120 pounds.

(b) The average income of persons was Rs 210 with S.D. of Rs 10 in a sample of 100 people of a city. For another sample of 150 person, the average income was Rs 220 with SD of Rs 12. Test whether there is any significant difference b/w the average incomes of localities

Soln We use Hypothesis Testing

Let $H_0 : \mu_1 - \mu_2 = 0$ (There is no significant difference in average incomes b/w two localities)

$H_1 : \mu_1 - \mu_2 \neq 0$ (There is a significant difference in average income b/w two localities)

Sample 1:

Mean $= \bar{x}_1 = Rs\ 210$

$S.D. = s_1 = Rs\ 10$

$n_1 = 100$

Sample 2:

$\bar{x}_2 = Rs\ 220$

$s_2 = Rs\ 12$

$n_2 = 150$

we use T stastic to solve

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}}$$

$$t = \frac{210 - 220}{\sqrt{\frac{(10)^2}{100} + \frac{(12)^2}{150}}}$$

$$t = \frac{-10}{\sqrt{1 + 0.96}}$$

$$t = \frac{-10}{\sqrt{1.96}} = -7.14$$

At 0.05 level of significance for two tailed test
The degrees of freedom are

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1}\right)}{n_1 - 1} \cdot \frac{\left(\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}\right)}{n_2 - 1}} = 109.23$$

So, the df at 0.05 level of significance comes out be approximately 109.23 and calculated value of t is -7.14.

Since the calculated value is greater than the tabular value (± 1.984) so we reject the null hypothesis at 5% level of significance. which suggests there is a significance difference between average incomes of two localities.

Q8 (a) The sales in supermarket during a week are given below. Test the hypothesis that the sales do not depend on the day of the week, using a significant level of 0.05.

Days	Mon	Tue	Wed	Thur	Fri	Sat	Total
Sales (Rs)	65	54	60	56	71	84	390

(b) Discuss various properties of χ^2-Test

Soln

Let's set up hypothesis

H_0 : The sales do not depend on the day of the week

H_1 : The sales depend on the day of the week.

· The Expected Values (E_i) are given by

$$E_{ij} = \frac{(\text{Total row } i) \times (\text{Total Column } j)}{\text{Grand Total}}$$

$$E_{11} = \frac{65 \times 390}{390} = 65 \qquad E_{13} = \frac{390 \times 60}{390} = 60$$

$$E_{14} = \frac{390 \times 56}{390} = 56$$

$$E_{12} = \frac{54 \times 390}{390} = 65$$

So Calculating so on the Expected frequency of each day is 65.

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

$$= \frac{(65-65)^2}{65} + \frac{(54-65)^2}{65} + \frac{(60-65)^2}{65} + \frac{(56-65)^2}{65} + \frac{(71-65)^2}{65} + \frac{(84-65)^2}{65} = \frac{121}{65} + \frac{25}{65} + \frac{81}{65} + \frac{36}{65} + \frac{361}{65}$$

$$= \underline{9.6}$$

Degrees of freedom $= (R-1) \times (C-1) = (6-1)(1-1) = 5$

Calculated Value of χ^2 stastistic at 5% level of significance is 9.6 at 5 d.o.f.

Tabular value is 11.07

Since Calculate value χ^2 (9.6) < Tabular χ^2 (11.07)

So we accept Null Hypothesis.

(b) Properties of Chi square statistics are:

(1) <u>Test for Independence</u>: Chi square Test is commonly used to test the independence of two categorical variables. It checks the observed frequencies and expected values.

(2) <u>Categorical variables</u>: The Chi square Test is applicable when dealing with categorical variables and not suitable for continuous variables.

(3) It Test significant association between Two categorical values.

(4) This Test assumes that the observed frequencies are independent

(5) The degrees of freedom (df) for Chisquare test depend upon number of categories being analysed

$df = (r-1) \times (c-1)$, r - no of rows, c - no of coloumns.

(6) Chi square stastistic: (χ^2) is calculated using

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

where $O_i \rightarrow$ observed frequency

$E_i \rightarrow$ Expected frequency

(7) If Calculated χ^2 is greater than critical value then Null hypotesis is Rejected. Indicating a significant association between the variables

x—x—x