Enrollment No.: _____

Darshan Institute of Engineering & Technology
B.Tech. | Sem-4 | Summer-2024

Course Code	:	2101HS402	**Date**	:	08-04-2024
Course Name	:	Probability & Statistics	**Duration**	:	150 Minutes
			Total Marks	:	70

Instructions:
1. Attempt all the questions.
2. Figures to the right indicates maximum marks.
3. Make suitable assumptions wherever necessary.

Q.1 **(A)** Probability distribution of a random variable X is given below. **4**

X	0	1	2	3	4
$P(X = x)$	0.1	K	0.5	0.2	0.1

Find $V(X)$.

(B) A problem is given to three persons P, Q, R whose respective chances of solving it are $\frac{2}{7}, \frac{4}{7}, \frac{4}{9}$ respectively. What is the probability that the problem is solved? **3**

OR

A bag contains 2 red, 3 green and 2 blue balls. Two balls are drawn at random. What is the probability that none of the balls drawn is blue?

(C) Three urns contain 6 red, 4 black; 4 red, 6 black, and 5 red, 5 black balls respectively. One of the urns is selected at random and a ball is drawn from it. If the ball drawn is red, find the probability that it is drawn from (i) the first urn, (ii) second urn, (iii) third urn. **7**

OR

Check whether the given function is probability function or not?
$f(x) = \begin{cases} |x| & ; -1 < x < 1 \\ 0 & ; \text{ otherwise} \end{cases}$. if yes then find variance.

Q.2 **(A)** A discrete random variable has probability mass function is given below: **4**

X	-2	-1	0	1	2	3
$P(X = x)$	0.2	k	0.1	2k	0.1	2k

Find k, $P(X \geq 1)$, $P(-1 < X \leq 2)$, $P(X < 0)$.

(B) A box of candies has many different colors in it. There is a 6% chance of getting a pink candy. What is the probability that at least 3 candies in a box are pink out of 100? **3**

OR

A random variable X has an exponential distribution with probability distribution function is given by $f(x) = \begin{cases} \frac{1}{5}e^{-\frac{x}{5}} & ; x > 0 \\ 0 & ; \text{ otherwise} \end{cases}$. then find probability that (a) X is not less than 3, (b) X < 3.

(C) The marks of 1000 students of university are found to be normally distributed with mean 70 and SD 5. Estimate the number of students whose marks will be (i) between 60 and 75, (ii) more than 75, (iii) less than 68. **7**

$P(0 < Z < 2) = 0.4772,\ P(0 < Z < 1) = 0.3413,\ P(0 < Z < 0.4) = 0.1554$

OR

A local report stated that the mean score on a placement test was 480 and 20% of the candidates scored below 400. Assume that the scores follow a normal distribution. Then find standard deviation of the scores and find percentage of the candidates that scored above 500.

Q.3 **(A)** The mean of the following distribution is 43. Find the value of p and mode of the given data. **4**

x_i	$20-30$	$30-40$	$40-50$	$50-60$	$60-70$
f_i	8	15	p	9	6

(B) Find the standard deviation of the first 7 prime numbers. **3**

OR

For a group of 10 items, $\sum x = 452,\ \sum x^2 = 24270$ and mode = 43.7. Find Karl Pearson's coefficient of Skewness.

(C) Find mean deviation about the mean, median and mode for the following data: **7**

x	2	5	6	8	10	12
f	2	8	10	7	8	5

OR

Calculate β_1 and β_2 for the following data:

x	1	2	3	4	5
f	2	3	5	4	1

Q.4 **(A)** A survey claims that 9 out of 10 doctors recommend aspirin for their patients with headaches. To test the claim, a random sample of 100 doctors is taken, and it is found that 85 of them recommend aspirin for their patients with headaches. At a significance level of 0.05, can we conclude that the proportion of doctors who recommend aspirin for their patients with headaches is different from 0.9? $(|Z_{0.05}| = 1.96)$ **4**

(B) Write Normal equation of the following curve: **3**
(i) $y = a + bx$, (ii) $y = a + bx + cx^2$, (iii) $y = ae^{bx}$

OR

Write the equation of Regression line (i) y on x (ii) x on y. Also, describe each term of the equation.

(C) The following data is obtained to know the relationship between maximum day temperature and the sale of ice-cream in Ahmedabad city. **7**

Temperature (Celsius)	35	42	40	39	44	40	45	40
Sale of ice cream (kg)	600	680	750	630	920	750	900	720

Calculate the rank correlation coefficient.

OR

Answer the following questions:

(i) Fit a parabola $y = a + bx + cx^2$ for the following data:

x	1	2	3	4	5
y	5	12	26	60	97

(ii) Fit a curve $y = ae^{bx}$ for the following data:

x	1	2	3	4
y	7	11	17	27

Q.5 **(A)** A researcher claims that the average time spent on social media by teenagers is less than 3 hours per day. A random sample of 20 teenagers is selected, and their daily social media usage times are recorded. The sample mean time spent on social media is found to be 2.5 hours with a standard deviation of 0.8 hours. Test the researcher's claim at a significance level of 0.01. $\left(t_{(0.01,\ 19)} = 2.540\right)$ **4**

(B) A random sample of fifteen paired observations from a bivariate population gives a correlation coefficient of -0.5. does this signify the existence of correlation in the sample population? $(t_{(0.05,\ 13)} = 2.1604)$ **3**

OR

The height of 8 males of a given locality are found to be $168, 175, 155, 170, 160, 165, 152, 170$ cm. Based on this sample, find the 95% confidence limit for the heights of males in the locality. $(t_{(0.05,\ 7)} = 2.365)$

(C) A researcher wants to investigate whether there is a significant difference in the preferences for three types of food among a group of 100 individuals. The observed frequencies are as follows: **7**

- Type A: 30 individuals
- Type B: 40 individuals
- Type C: 30 individuals

Conduct a chi-square test of goodness of fit to determine if there is a significant difference in the preferences for these three types of food. Use a significance level of 0.05. Provide all necessary calculations and interpret the results.
$(\chi^2_{0.05}(v = 2) = 5.99)$

OR

Two nicotine contents in two random samples of tobacco are given below:

Sample − I	21	24	25	26	27	−
Sample − II	22	27	28	30	31	36

Can we say that two samples came from the same normal population?

$$(F_{0.05}(5, 4) = 5.19)\ \left(t_{(0.05,\ 9)} = 2.26\right)$$

***