

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
**MCA– SEMESTER -II EXAMINATION- SUMMER-2023**

**Subject Code: 629409****Date: 14/07/2023****Subject Name: Statistical Methods****Time: 10:30 AM TO 01:00 PM****Total Marks: 70****Instructions:**

1. Attempt all questions.
2. Make Suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Use of simple calculators and non-programmable scientific calculators are permitted.

- Q.1 (a)** Classify each of the following as nominal, ordinal, interval or data. **07**
1. The time required to produce each tire on an assembly line
  2. The number of quarts of milk a family drinks in a month
  3. The ranking of four machines in your plant after they have been designated as excellent, good, satisfactory, and poor
  4. The telephone area code of clients in the United States
  5. The age of each of your employees
  6. The dollar sales at the local pizza shop each month
  7. An employee's identification number

- (b)** The owner of a fast-food restaurant ascertains the ages of a sample of customers. From these data, the owner constructs the frequency distribution shown. For each class interval of the frequency distribution, determine the class midpoint, the relative frequency, and the cumulative frequency. **07**

Class Interval	Frequency
0–under 5	6
5–under 10	8
10–under 15	17
15–under 20	23
20–under 25	18
25–under 30	10
30–under 35	4

What does the relative frequency tell the fast-food restaurant owner about customer ages?

- Q.2 (a)** According to the National Retail Federation and Center for Retailing Education at the University of Florida, the four main sources of inventory shrinkage are employee theft, shoplifting, administrative error, and vendor fraud. The estimated annual dollar amount in shrinkage (\$ millions) associated with each of these sources follows: **07**

Employee theft            \$17,918.6

Shoplifting                    15,191.9

Administrative error    7,617.6

Vendor fraud                2,553.6

---

Total                            \$43,281.7

Construct a pie chart and a bar chart to depict these data.

- (b) (i) A company has 140 employees, of which 30 are supervisors. Eighty of the employees are married, and 20% of the married employees are supervisors. If a company employee is randomly selected, what is the probability that the employee is married and is a supervisor? 07
- (ii) A manufacturing firm produces pads of bound paper. Three percent of all paper pads produced are improperly bound. An inspector randomly samples two pads of paper, one at a time. Because a large number of pads are being produced during the inspection, the sampling being done, in essence, is with replacement. What is the probability that the two pads selected are both improperly bound?

**OR**

- (b) Compute the, mean, median, mode ,variance and standard deviation on the following sample data: 07

Class Interval	Frequency	Cumulative Frequency
10–under 15	06	06
15–under 20	22	28
20–under 25	35	63
25–under 30	29	92
30–under 35	16	108
35–under 40	08	116
40–under 45	04	120
45-under 50	02	122

- Q.3** (a) Machines A, B, and C all produce the same two parts, X and Y. Of all the parts produced, machine A produces 60%, machine B produces 30%, and machine C produces 10%. In addition, 40% of the parts made by machine A are part X. 50% of the parts made by machine B are part X. 70% of the parts made by machine C are part X. A part produced by this company is randomly sampled and is determined to be an X part. With the knowledge that it is an X part, revise the probabilities that the part came from machine A, B, or C. 07

- (b) Bank customers arrive randomly on weekday afternoons at an average of 3.2 customers every 4 minutes. What is the probability of having more than 7 customers in a 4-minute interval on a weekday afternoon? 07

**OR**

- Q.3** (a) What is the probability of obtaining a score greater than 700 on a GMAT test that has a mean of 494 and a standard deviation of 100? Assume GMAT scores are normally distributed. 07

$$P(X > 700/\mu = 540 \text{ and } \sigma = 100) = ?$$

- (b) (i) A manufacturing firm has been involved in statistical quality control for several years. As part of the production process, parts are randomly selected and tested. From the records of these tests, it has been established that a defective part occurs in a pattern that is Poisson distributed on the average of 1.38 defects every 20 minutes during production runs. Use this information to determine the probability that less than 15 minutes will elapse between any two defects. 07
- (ii) Suppose that during any hour in a large department store, the average number of shoppers is 448, with a standard deviation of 21 shoppers. What is the probability that a random sample of 49 different shopping hours will yield a sample mean between 441 and 446 shoppers?

- Q.4 (a)** A study is conducted in a company that employs 800 engineers. A random sample of 50 engineers reveals that the average sample age is 34.3 years. Historically, the population standard deviation of the age of the company's engineers is approximately 8 years. Construct a 98% confidence interval to estimate the average age of all the engineers in this company. **07**
- (b)** A survey of the morning beverage market shows that the primary breakfast beverage for 17% of Americans is milk. A milk producer in Wisconsin, where milk is plentiful, believes the figure is higher for Wisconsin. To test this idea, she contacts a random sample of 550 Wisconsin residents and asks which primary beverage they consumed for breakfast that day. Suppose 115 replied that milk was the primary beverage. Using a level of significance of .05, test the idea that the milk figure is higher for Wisconsin. **07**

**OR**

- Q.4 (a)** Test the following hypotheses of the difference in population means by using the following data ( $\alpha = 0.10$ ) and the eight-step process. **07**

$$H_0: \mu_1 - \mu_2 = 0 \quad H_a: \mu_1 - \mu_2 < 0$$

Sample-1	Sample-2
$\bar{x}_1 = 51.3$	$\bar{x}_2 = 53.2$
$\sigma_1^2 = 52$	$\sigma_2^2 = 60$
$n_1 = 31$	$n_2 = 32$

- (b)** Sketch a scatter plot from the following data, and determine the equation of the regression line. **07**

x	140	119	103	91	65	29	24
y	25	29	46	70	88	112	128

- Q.5 (a)** Test the slope of the regression model developed in Demonstration Problem 12.1 to predict the number of FTEs in a hospital from the number of beds to determine whether there is a significant positive slope. Use  $\alpha = .01$  **07**
- (b)** Compute the coefficient of determination  $r^2$  in which a regression model was developed to predict the number of FTEs of a hospital by the number of beds. **07**

Number of Beds	FTEs	Number of Beds	FTEs
23	69	50	138
29	95	54	178
29	102	64	156
35	118	66	184
42	126	76	176
46	125	78	225

**OR**

- Q.5 (a)** Define with example **07**
1. SSE
  2. Standard error of estimate
  3. Coefficient of Multiple Determination ( $R^2$ )
  4. Adjusted  $R^2$
- (b)** The client company data from the Decision Dilemma reveal that 155 employees worked one of four types of positions. Shown here again is the raw values matrix (also called a contingency table) with the frequency counts for each category and for subtotals and totals containing a breakdown of these employees by type of position and by sex. If an employee of the company is selected randomly, what is the probability that the employee is female or a professional worker? **07**

\*\*\*\*\*