Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

MCA INTEGRATED– SEMESTER -III EXAMINATION –WINTER-2022

Subject Code: 2638602 Date: 29/12/2022

Subject Name: Basic Statistics

Time: 02:30 PM TO 5:00 PM Total Marks: 70

Instructions:

1. **Attempt all questions.**
2. **Make Suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**

Q.1 **(a)** Define the following terms: **07**

1) Skewness
2) Mean Absolute Deviation
3) Kurtosis
4) Null hypothesis
5) Marginal Probability
6) Mutually Exclusive Event
7) Chebyshev's Theorem

(b) Find the median of the following frequency distribution: **07**

Weekly wages (in Rs.)	60-69	70-79	80-89	90-99	100-109	110-119
No of Days:	5	15	20	30	20	8

Q.2 **(a)** A bag contains 8 red, 3 white and 9 blue balls. If three balls are drawn at random, determine the probability that **07**
(i) all the three balls are blue balls
(ii) all the balls are of different colours.

(b) State the Baye's rule .In a bolt factory, machines A, B and C manufacture 25%, 35%, 40% respectively. Of the total of their output 5, 4 and 2% are defective. A bolt is drawn and is found to be defective. What are the probabilities that it was manufactured by the machines A ? **07**

OR

(b) Find variance for following data: **07**

Class interval	30−40	40−50	50−60	60−70	70−80	80−90	90−100
Frequency	3	7	12	15	8	3	2

Q.3 **(a)** What is correlation? Calculate the correlation coefficient between x and y for the following data: **07**

x	5	9	13	17	21
y	12	20	25	33	35

(b) What is Binomial distribution? A student takes a ten-question true/false exam. **07**

a. Find the probability that the student gets exactly six of the questions right simply by guessing the answer on every question.
b. Find the probability that the student will obtain a passing grade of 60% or greater simply by guessing.

OR

Q.3 **(a)** Find the inter quartile range for the data: 9,11,15,19,17,13,7. **07**

(b) List all types of sampling methods. Explain any two Probabilistic and any two non-probabilistic sampling methods. **07**

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) The average number of annual trips per family to amusement parks is Poisson distributed,with a mean 0.7 trips per year. **07**

1. What is the probability that family took exactly one trip to an amusement park last year?
2. The family took three or fewer trips to amusement parks over a three-year period.

OR

Q.4 **(a)** Give the difference Between (i) One tailed and two tailed test (ii) Type-I and Type II Error. **07**

(b) The average number of acres burned by forest and range fires in a large New Mexico county is 4,300 acres per year, with a standard deviation of 750 acres. The distribution of the number of acres burned is normal. What is the probability that between 2,500 and 4,200 acres will be burned in any given year? **07**

Q.5 **(a)** Develop the equation of the simple regression line y on x for the following data: **07**

X	10	9	8	7	6	5
Y	8	12	7	10	9	6

(b) A small business has 37 employees. Because of the uncertain demand for its product, the company usually pays overtime on any given week. The company assumed that about 50 total hours of overtime per week is required and that the variance on this figure is about 25. Company officials want to know whether the variance of overtime hours has changed. Given here is a sample of 16 weeks of overtime data (in hours per week). Assume hours of overtime are normally distributed. Use these data to test the null hypothesis that the variance of overtime data is 25. Let α = .10. **07**

57 56 52 44 46 53 44 44 48 51 55 48 63 53 51 50

OR

Q.5 **(a)** Determine the equation of the regression line for the following data, and compute the residuals. **07**

X	15	8	19	12	5
Y	47	36	56	44	21

(b) According to the U.S Bureau of labor statistics, the average weekly earnings of a production worker in 1997 were $ 424.20. Suppose a labor researcher wants to test to determine whether this figure is still accurate today. The researcher randomly selects 54 production workers from across the united states and obtains a representative earnings statement for one week from each. The resulting sample average is $ 432.69. Assuming a population standard deviation of $ 33.90, and a 5% level of significance, test the hypothesis that the mean weekly earnings of a production worker have changed. **07**

****