Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

MCA INTEGRATED– SEMESTER-III EXAMINATION –SUMMER-2020

Subject Code: 2638602 Date: 05-11-2020

Subject Name: Basic Statistics (BS)

Time: 02:30 PM to 05:00 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 briefly the following terms: **07**

i) Median
ii) Bar chart
iii) Mutually exhaustive events
iv) Snowball sampling method
v) Skewness
vi) Correlation
vii) Level of Significance

(b) Consider a sample with data values 37, 20, 35, 22, 15, 37, 29 and 26. Compute Five number summary. Also draw the box plot. **07**

Q.2 **(a)** For the following data obtain correlation coefficient. **07**

X	22	23	25	27	28	30	32	34	35	37
Y	11	14	15	17	13	19	22	24	26	29

(b) Given the probabilities of three events, A, B, and C, occurring are P(A) = 0.45, P(B) = 0.45, and P(C) = 0.1. Assuming that A, B, or C has occurred, the probabilities of another event, X, occurring are P(X/A) = 0.7, P(X/B) = 0.75, and P(X/C) = 0.4. Find P(A/X), P(B/X), and P(C/X). **07**

OR

(b) Discuss different types of Sampling methods with example. **07**

Q.3 **(a)** Consider a sample with data values of 27, 25, 20, 15, 30, 34, 28 and 25. Obtain Mean, Median, Mode, standard deviation and IQR. **07**

(b) A data firm records a large amount of data. Historically, .9% of the pages of data recorded by the firm contain errors. If 200 pages of data are randomly selected, **07**

a. What is the probability that six or more pages contain errors?
b. What is the probability that more than 10 pages contain errors?
c. What is the probability that none of the pages contain errors?
d. What is the probability that fewer than five pages contain errors?

OR

Q.3 **(a)** In a sample of 500 people from a village, 270 are found to be vegetarians and the rest are non-vegetatians. Can we assume that both vegetarians and non-vegetarians are equally popular? **07**

(b) Suppose 70% of all companies are classified as small companies and the rest as large companies. Suppose further, 82% of large companies provide training to employees, but only 18% of small companies provide training. A company is randomly selected without knowing if it is a large or small company; however, it is determined that the company provides training to employees.What are the prior probabilities that the company is a large company or a small company? What are the revised probabilities that the company is large or small? Based on your analysis, what is the overall percentage of companies that offer training? **07**

Q.4 **(a)** The U.S. Bureau of Labor Statistics reports that of persons who usually work full-time, the average number of hours worked per week is 43.4. Assume that the number of hours worked per week for those who usually work full-time is normally distributed. Suppose 12% of these workers work more than 48 hours. Based on this percentage, what is the standard deviation of number of hours worked per week for these workers? **07**

(b) The average cost of a one-bedroom apartment in a town is $550 per month. What is the probability of randomly selecting a sample of 50 one-bedroom apartments in this town and getting a sample mean of less than $530 if the population standard deviation is $100? **07**

OR

Q.4 **(a)** A random sample of size 39 is taken from a population of 200 members. The sample mean is 66 and the population standard deviation is 11. Construct a 96% confidence interval to estimate the population mean. What is the point estimate of the population mean? **07**

(b) Give Difference between (i) One tailed and two tailed test (ii) Type-I and Type-II Error. **07**

Q.5 **(a)** Suppose the following data are selected randomly from a population of normally distributed values. **07**

40, 51, 43, 48, 44, 57, 54, 39, 42, 48, 45, 39, 43

Construct a 95% confidence interval to estimate the population mean.

(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, determine whether the mean weekly earnings of a production worker have changed. **07**

OR

Q.5 **(a)** Given the data for two variables x and y. **07**

X	12	14	16	18	20
Y	25	28	30	35	44

a. Develop an estimated regression equation for these data.

b. Compute the residuals.

(b) A study by Hewitt Associates showed that 79% of companies offer employees flexible scheduling. Suppose a researcher believes that in accounting firms this figure is lower. The researcher randomly selects 415 accounting firms and through interviews determines that 303 of these firms have flexible scheduling. With a 1% level of significance, does the test show enough evidence to conclude that a significantly lower proportion of accounting firms offer employees flexible scheduling? **07**

****