

Seat No.: _____

Enrolment No. _____

GUJARAT TECHNOLOGICAL UNIVERSITY
M.Sc. Industrial Biotechnology/Postgraduate Diploma in Bioinformatics
SEMESTER - 1 SUMMER – 2023 EXAMINATION

Subject Code:1310105

Date:23 Jun 2023

Subject Name: Statistics

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.
4. Draw neat and clean diagrams as required

Q.1 Write a note on following

**(Marks-
10X2=20)**

1. Types of categorical data.
2. What does the histogram describe?
3. Difference between the Z-test and t-test.
4. What statistical test is used to compare more than two groups?
5. What are the types of non-random sampling?
6. Calculate the standard deviation and coefficient of variance from the following data
Variable= 10, 30, 50, 70, 20, 60, 40
7. What is the concept of linear regression?
8. Which measure central tendency is most informative when data is skewed? Give a brief explanation about it.
9. Calculate the median from the following data

Classes	Frequency	Cumulative frequency
0-10	7	
10-30		
30-60	28	
60-80	10	
80-90	8	
90-100	3	70

Q1. Find out the missing frequency

Q2 Find out the median

10. Difference between one-tailed and two-tailed testings.

Q.2 Answer the following (Any 2 out of 3)

**(Marks-
2X10=20)**

1. A training program was conducted to improve students' knowledge. Data were collected from some students both before and after the training program

Pre Training	Post Training
12	15
14	18
13	19
11	18
12	14
10	15
15	19
13	13
9	12
14	16

Find out the central tendency and variability of data and apply a suitable statistical test to find out whether the training program is effective or not (if $\alpha=0.05$)

2. Data on annual rainfall and the number of clusters per plant are noted on 12 randomly selected plant neem tree

Annual rainfall (cm)	No. of clusters
9	12
8	10
12	14
7	9
13	10
12	12
17	8
21	22
20	18
19	17

Calculate the regression coefficient (b) and its test of significance

3. Describe the principle of experimental design and explain different types of experimental design with their pros and cons.

Q.3

Answer the following (Any 6 out of 8)

- Describe the various methods of studying correlation.
- What are the different types of distribution? Explain briefly.
- A study was designed to test the IL-9 levels in patients with diabetes, diabetic Nephropathy and healthy control. Data are shown below

Healthy control (pg/ml)	Diabetic (pg/ml)	Diabetic Nephropathy (pg/ml)
152	550	989
220	685	1080
110	640	1182
257	596	1074
130	620	1221
200	548	1150

Find out if there is any statistical difference in IL-9 levels between the above groups.

4. Briefly describe the method sampling.

(Marks-6X5=30)

5. Describe the theorems of probability and applications of probability.
6. What are completely Randomized design and least square designs? Write down its advantages and disadvantages.
7. The table shows the scores for ten students in History and Geography exam papers.

History	54	89	65	91	74	99	53	45	50	85
Geography	86	66	72	59	82	61	79	69	74	80

Find the average score for the History and Geography papers.
 What kind of correlation exists between the two exam papers?

8. In a hospital, 300 cases of malaria were admitted at a particular time, 150 cases were given Chloroquine, and 150 were given Artemisinin. Find whether the difference is significant or not using chi-square test

Drug	Cured	Not cured	Total
Chloroquine	143	7	150
Artemisinin	137	13	150
Total	280	20	300
