M. Sc. (Biotechnology) Examination, 2023

(1st Year, 1st Semester)

Biostatistics and Computational Biology

Paper – CORE/TH/134

Time : Two hours   Full Marks : 40

Group – A

Answer any **three** questions.

1. Two microbial groups, Gram-negative and Gram-positive, consist of 100 bacterial strains each. The Gram-negative bacteria group was treated with a newly developed antibiotic A. The group Gram-positive (which is called the control group); was not heated. Otherwise, the two groups were treated identically. The servility result indicates 75 and 65 Gram-negative, respectively, survived after 6 hours of treatment.

    a) Explain how the data for this hypothesis test is appropriate for a Z test.

    b) Test the hypothesis that the Gram-negative bacterial strains are more sensitive to the tested antibiotic using the level of significance of i) 0.01 ($Z_{0.01}$ = 2.33) and ii) 0.05 ($Z_{0.05}$ = 1.64).

    c) Considering the results of the hypothesis test, decide which Type I or Type II errors are possible, and describe this error.   2+(4+2)+2

[Turn over

2. a) What are the characteristic features of Normal or Binomial distribution in statistics?

   b) The mean length of 500 different eukaryotic cells under microscopy is 151 nm and the Standard deviation is 15 nm. Assuming that the length is normally distributed, find how many cells measured (i) between 120 and 155 nm, (ii) more than 185 nm ($Z_{2.06} = 0.0192$, $Z_{0.27} = 0.3897$, $Z_{0.27} = 0.0116$)

   2+4+4

3. Voters in a particular state are surveyed. Each respondent is assigned an identification number, and information about each of the following is recorded: sex of the voter and age

Sex	Age
M	35
F	20
F	35
M	41
M	39
M	59
F	20
M	52
F	44
M	46
F	40
F	34
F	24
M	62
M	44

   b) Write the appropriate script that are used for the following object in R.

      1. Objects are currently defined in the working directory.

      2. Objects can be removed using what function?

      3. To clear the workspace

      4. To check the current working directory.

      2+(2+2+2+2)

3. A class of 10 students have been surveyed and their heights were recorded:

   150cm, 150cm, 142cm, 154cm, 168cm, 153cm, 151cm, 153cm, 142cm and 151cm

   a) How the values would be stored in a variable?

   b) How to recall a single value from the vector?

   c) Use the function for calculating the i) Five-number summary, ii) Variance, iii) Standard deviation and iii) Histogram.

   2+2+6

a) Complete the construction of the figure to obtain a stem-and-leaf display for males and a stem-and-leaf display for females.

b) Find the five-number summary for males and the five-number summary for females.

c) Construct a box plot for males and a box plot for females.   4+(2+2)+2

4. The survey on random people from two different districts of West Bengal, Kolkata and North 24 Parganas is asked about their willingness to take the vaccine. The results of the survey are displayed in the contingency table below.

	Willingness to take the vaccine	
Residence	YES	NO
Kolkata	127	81
North 24 Parganas	248	344

a) Explain how the data for this hypothesis test is appropriate for a chi-square test concerning independence.

b) A 0.05 significance level is chosen for a hypothesis test to see if there is any evidence of a relationship between locations of residence and whether or not willing to take the vaccine ($\chi^2_{1;\ 0.05} = 3.841$).

c) Considering the results of the hypothesis test, decide which Type I or Type II errors are possible, and why.

1+8+1

---

5. A private school principal claims that the students in his class are above average in intelligence. A random sample of nine students' IQ scores are 89, 99, 105, 116, 116, 118, 119, 125 and 128. Is there sufficient evidence to support the principal's claim? The average IQ of all the students of this school is 100 with a standard deviation of 15. (Given $t_{8;\ 0.05} = 1.860$ and $t_{8;\ 0.01} = 2.896$)

a) Explain how the data for this hypothesis test is appropriate for a Student's t-test.

b) Decide whether H0 would have been rejected or accepted with the following significance level: (i) $\alpha = 0.01$, (ii) $\alpha = 0.05$.

c) Considering the hypothesis test result, decide which Type I or Type II errors are possible and explain.

2+(4+2)+2

Group – B

Answer any **one** question.

1. How to set a working director in R? How can you load a .csv file in R? How do you install a package in R? Differentiate between Vector, List, Matrix, and Data frame.   2+2+2+4

2. a) Construct a matrix with 3 rows that contain the numbers 1 up to 15?