

Date:12-03-2022

Registration number:

ST. JOSEPH’S COLLEGE (AUTONOMOUS), BENGALURU-27

M.Sc. MICROBIOLOGY - III SEMESTER

SEMESTER EXAMINATION: OCTOBER 2021

(Examination conducted in January-March 2022)

**MB 9418 Biostatistics and Bioinformatics**

Time- 2 ½ hrs Max Marks-70

**Answer section A and section B in two separate answer booklets.**

**SECTION A – BIOSTATISTICS**

**I. Answer any FIVE of the following questions 5x2=10**

1. How do you expect the graph to look like when mode is larger than mean?
2. What is the difference between causation and correlation?
3. What is the significance of Mann-Whitney U test?
4. State two differences t - test and chi-square test.
5. The following table gives the weight of 31 microbial colonies in micrograms in a sample enquiry. Calculate the mean weight.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Weight (X) | 130 | 135 | 140 | 145 | 146 | 148 | 149 | 150 | 157 |
| No. of colonies (f) | 3 | 4 | 6 | 6 | 3 | 5 | 2 | 1 | 1 |

1. If the random variable X follows a Poisson distribution with mean 3.4, find P(X=6).
2. Why do we calculate p-value?

**II. Answer any TWO of the following questions 2x5 =10**

1. What do you understand by confidence interval of an estimate? What is its significance? Explain using suitable example.
2. What is the significance of statistics in the field of microbiology? Explain using suitable example.
3. An anti-bacterial drug test is known to produce 10% false results. If the random variable X is the number of false tests produced in run of 4 tests, find the probability that X takes the value of 0 to 4.

**III. Answer any ONE of the following questions 1x10=10**

11. Determine the equation for the line of best fit from the data given below and calculate the value of variable:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | 8 | 4 | 5 | -1 |
| Y | -2 | 0 | 2 | 6 |

12. What is Randomization? What are the different ways in which randomization can be achieved? Explain one of the methods of randomization in detail using an example.

**IV. Answer the following question 1x5=5**

13. Calculate Karl Pearson’s correlation between X and Y series from data given below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **X series** | 12 | 9 | 8 | 10 | 11 | 13 | 7 |
| **Y series** | 14 | 8 | 6 | 9 | 11 | 12 | 3 |

**SECTION B - BIOINFORMATICS**

**Answer section B in a separate answer booklet.**

**I. Answer any FIVE of the following 5x2=10**

1. What is a reference pathway in KEGG database?
2. Describe any two broad types of restriction enzymes that can be retrieved using NEBCutter.
3. Write a brief note on any two NCBI resources.
4. Write any two advantages of NGS technology over Sanger method in microbiology.
5. Write a note on query designing for literature search.
6. List any six categories of biological information in Entrez Gene.
7. Write a note on FASTA format.

**II.** **Answer any TWO of the following 2x5= 10**

1. Briefly describe the prediction methods used in STRING protein interaction resource.
2. List any five bioinformatics resources and their uses in microbiology research.
3. Write a note on different types of BLAST programs along with their specific applications.

**III. Answer any ONE of the following 1x10=10**

1. Differentiate between Roche and Illumina technologies. How is NGS advantageous over Sanger method?
2. List any three differences between BLAST and ClustalOmega. Explain the output features of ClustalOmega.

**IV. Answer the following 1x5=5**

1. Describe various approaches in metagenomics. List its applications in microbiology field.