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

Registration Number:

Date & Session:9-12-22

**M.Sc (MICROBIOLOGY) – III SEMESTER**

**SEMESTER EXAMINATION: OCTOBER 2022**

**(Examination conducted in December 2022)**

**MB 9421: BIOSTATISTICS AND BIOINFORMATICS**

**Time: 2 ½ Hours Max Marks: 70**

**This paper contains 2 printed pages and 4 parts**

1. **Answer any Five of the following 5X3=15**

1. Differentiate between median and mode.
2. State the significance of goodness of fit.
3. Define regression.
4. What are ORFs? What is their significance?
5. Write the salient features of next-gen sequencing.
6. Explain the need of hidden Markov models in statistics.
7. How can we analyse the results of mass fingerprinting?
8. **Answer any Five of the following 5X5=25**
9. Draw an ogive to determine the median of the following data.

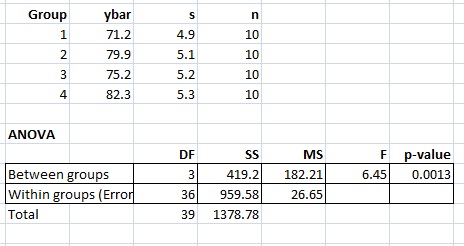
|  |  |
| --- | --- |
| Class interval | frequency |
| 0-10 | 10 |
| 10-20 | 15 |
| 20-30 | 18 |
| 30-40 | 25 |
| 40-50 | 13 |
| 50-60 | 9 |
| 60-70 | 5 |

1. Write a short note on EMBL database.
2. Define following terms:
   1. p-value
   2. Confidence interval
   3. Level of significance
3. What is the significance of hypotheses and hypotheses testing?
4. A random collection of 100 water samples collected across four major Indian water basins to test whether microbial species richness varies across the different water basins. The data are shown in the table below. Enlist your null hypothesis and analyse and interpret your result. (Note for d.f. = 3, significance level = 0.01, ϰ2=11.341)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Major River Basins of India | | | | Total |
|  | Brahmaputra | Cauvery | Godavari | Indus |
| No. of species | 30 | 40 | 10 | 20 | 100 |

1. Enumerate the major outcomes of human genome project.
2. Explain the process of analysing and interpreting the results of HPLC.
3. **Answer any Two of the following 2X10=20**
4. Describe the salient features of maximum parsimony approach.
5. Discuss the role of PDB in the field of bioinformatics.
6. Explain the key aspects to consider while making a phylogenetic tree using MEGA software.
7. **Answer the following 1X10=10**
8. Calculate the Least Significant Difference for the difference between two means on Group 1 and Group 2 with the following test results:

Given: The t-critical value is 2.028 (for α = 0.05, dfw = 36)

[](https://www.statisticshowto.com/wp-content/uploads/2013/11/lsd.jpg)