

Question Paper

Exam Date & Time: 27-Dec-2023 (02:00 PM - 05:00 PM)



MANIPAL ACADEMY OF HIGHER EDUCATION

THIRD SEMESTER M.Sc. (MEDICAL BIOTECHNOLOGY / MOLECULAR BIOLOGY AND HUMAN GENETICS / GENOME ENGINEERING / TISSUE ENGINEERING) DEGREE EXAMINATION - DEC 2023 / JAN 2024
SUBJECT: MBT601 / MBH601 / MGE 603 / MTE 603 BIOSTATISTICS AND BIOINFORMATICS
(OBE-2021 REVISED REGULATION - REGULARS)

Marks: 70

Duration: 180 mins.

SECTION A - BIOSTATISTICS (35 MARKS)

Answer all the questions.

- 1A) Identify the type of variable (Nominal/Ordinal/Discrete/Continuous) (3)
i) Gender
ii) Lifespan of fruitflies (in completed days)
iii) Residual sugar content (g/l) in wine
- 1B) Define the following terms used in statistical inference: (2)
i) Type II error
ii) Sampling distribution
- 1C) Gene mutations have been found in patients with muscular dystrophy. In a study, it was found that there were defects in the gene coding of sarcoglycan proteins in 23 of 184 patients with limb-girdle muscular dystrophy. Use these data to construct a 90% confidence interval for the corresponding population proportion. (4)
- 1D) Distinguish between case-control study and a cohort study. (5)
- 2A) A certain form of cancer is known to be found in women over 60 years of age with probability 0.08. A blood test exists for the detection of the disease but the test is not infallible. In fact, it is known that 2% of the time the test gives a false negative (i.e., the test incorrectly gives a negative result) and 10% of the time the test, gives a false positive (i.e., incorrectly gives a positive result). If a woman over 60 years is known to have taken the test and found negative for cancer, what is the probability that she has the disease? (3)
- 2B) A certain drug treatment cures 80% of cases of hookworm in children. Suppose that 12 children suffering from hookworm are to be treated, and that the children can be regarded as a random sample from the population. Find the probability that:
i) All but one will be cured
ii) At least one will be cured (4)
- 3) In a study of the lizard *Sceloporus occidentalis*, biologists measured the distance (m) run in two minutes for each of 14 animals. The results (listed in increasing order) were as follows:
18.4 22.2 24.5 26.4 27.5 28.7 30.6 32.9 32.9 34.0 34.8 37.5 42.1 45.5
Compute quartile deviation. (7)
- 4) Distinguish between linear and logistic regression. (3.5)
- 5) A study was planned to find whether there is any difference in the average RBC Cholinesterase values (measured in micro mol/min/ml) between alcoholic and non-alcoholic adult males. What should be the minimum sample size required in each group to detect a clinically significant difference of 3 micro mol/min/ml at 90% power and 5% level of significance? Assume the pooled standard deviation of RBC Cholinesterase values is 5 micro mol/min/ml. ($Z_{1-\alpha/2} = 1.96$, $Z_{1-\beta} = 1.28$) (3.5)

PART B - BIOINFORMATICS (35 MARKS)

Answer all the questions.

Illustrate where necessary.

- 6A) Describe the workflow employed for the analysis of microarray data. (3.5)
- 6B) Write a note on docking studies. (3.5)
- 7A) How does PubMed process its queries? (7)
- 7B) What are different types of BLAST? Add a note on the score and E-Value used in interpreting BLAST results. (7)
- 8) What are the steps to be followed for building a phylogenetic tree? Explain in detail the distance based method of tree building. (14)

-----End-----