MAM-003-038305   Seat No. _____

B. Voc. (Medical Laboratory & Molecular Diagnostic Technology) (Sem. III) Examination

October / November – 2016

GMLMDT - 3.5 : Introduction to Bioinformatics & Biostatistics

Faculty Code : 003
Subject Code : 038305

Time : $2\frac{1}{2}$ Hours]   [Total Marks : **70**

Instructions :
    (1)   All questions are compulsory.
    (2)   Figures on right indicate marks.

1   Answer the following :   **20×1=20**

    (1)   _____ is a collection of procedures and principles for gathering data and analyzing information to help people make decisions when faced with uncertainty.

    (2)   _____ referring to numbers or nonnumerical labels (male/female) collected from a set of entities (people, cities, etc).

    (3)   Median of a numerical list of data is the value _____.

    (4)   _____ are observations (such as measurements, genders, survey responses) that have been collected.

    (5)   _____ are collected when all individuals in a population are measured.

    (6)   The branch of biostatistics which deals with methods of collection, organization and presentation of Data is called

    (7)   The characteristic or quantity that may vary from one individual to another is called?

    (8)   Write down the name of two parametric test.

    (9)   The derived databases are also known as

    (10) Term Bioinformatics was coined by

    (11) The information retrieval tool for the NCBI is

    (12) Which of the following statistics measures the most frequently occurring value in a set of data?

    (13) Measures of lack of symmetry is called

(14) The type of sampling in which each element of population has equally likely chance of occurrence in a random sample is classified as

(15) Who is regarded as father of Biostatistics?

(16) Variable whose value can be expressed numerically is called as

(17) The degree of relation between two variables is studied through the concept of

(18) If correlation coefficient between X and Y is equal to –1 the variable are known as

(19) Degree of Freedom for students t test when n=20 and n=12 are

(20) Normal distribution is _____ curve.

2 (a) Answer in brief : (any **3**)   **3×2=6**

    (1) Explain how to draw frequency curve.

    (2) Give an account of the types of correlation.

    (3) What is null hypothesis ?

    (4) Give an account on the regression analysis.

    (5) What is a database? Discuss types of database.

    (6) Give an account of the procedure for testing of hypothesis.

(b) Answer in brief : (any **3**)   **3×3=9**

    (1) Discuss the types of available software for Bioinformatics.

    (2) Find the median score of the following scores obtained by students in a particular one hour examination.
6, 9, 13, 4, 11, 8, 12, 9, 5, 7

    (3) Find the range of for the following weight of 10 goats
30, 25, 15, 13, 16, 17, 20, 12, 50, 60

    (4) Define Skewness.

    (5) Give the name of retrieval tools for DDBJ, EMBL and Gen Bank respectively.

    (6) The following table lists the number of days that five houses had been up for sale, as well as their selling price, calculate the mean

X (Days)	45	12	3	17	32
Y	275	401	420	212	365

(c) Answer in detail : (any **2**)   **2×5=10**

   (1) Write down short note on INSDC

   (2) Write down the types of correlation

   (3) Differentiate between primary and secondary data

   (4) Explain various steps involved in T test.

   (5) Distinguish between primary data and secondary data by giving suitable example.

3 (a) Answer in brief : (any **3**)   **3×2=6**

   (1) Describe advantages of random sampling.

   (2) Discuss various steps in the procedure of testing the hypothesis

   (3) What do you mean by presentation of data? Describe various methods of presenting data collected by investigators.

   (4) Explain the need for Analysis of Variance

   (5) Write short note on kurtosis

   (6) Define variance.

   (b) Answer in brief : (any **3**)   **3×3=9**

   (1) Find out mean, median and mode for the distribution of the weights of 150 sheep from the data given below :

Weight in kg	05 − 10	10 − 15	15 − 20	20 − 25	25 − 30
Frequency	18	37	45	27	15

   (2) Discuss some common biological file formats.

   (3) Discuss modern Drug Discovery in detail

   (4) Write short note on FASTA

   (5) Write a short note on various measure of variance

   (6) Write short note on OMICS technology

(c) Answer in detail : (any **2**) **2×5=10**

(1) The following are the hemoglobin values (g/100 ml) of 10 children receiving a treatment for hemolytic anemia : 8.1, 9.9, 10, 9.5, 8.5, 9.1, 12.4, 7.5, 9.8, 6.7 Compute mean, median and standard deviation.

(2) 6 children were given mother's milk while the second groups of 5 children of same age group were given dairy milk. After 1 year the gain in weight in kg. was noted. The noted weight in both groups is given below (Apply t test to find out significant difference between means of two groups at 5% level of significance.) $t_{0.05, 9} = 2.26$

X_1	4	3	4	3	2	6
X_2	1	3	2	4	4	

(3) Find whether the following phenotype distribution in a sample of 192 flies from a particular fruit fly population has a goodness of fit with the Mendalian ratio 9 : 3 : 3 : 1 $\left(\chi^2_{0.05, 3} = 7.82 \right)$

Grey body Red Eyed	Black body Red Eye	Grey Body Scarlet Eye	Black body Scarlet Eye
100	52	20	20

(4) Calculate the correlation co-efficient between X and Y from the following data :

X	6	9	15	4	21
Y	12	20	25	33	35

(5) From the data given below, find out whether the means of the three samples differ significantly or not. $F_{0.05, 2, 12} = 3.9$

Sample 1	20	10	17	17	16
Sample 2	19	13	17	12	9
Sample 3	13	12	10	15	5