

Exam. Code : 206601

Subject Code : 5122

M.Sc. Bio-Informatics 1<sup>st</sup> Semester

BASIC BIOSTATISTICS

Paper—BI-513

Time Allowed—Three Hours] [Maximum Marks—75

**Note** :— Attempt **SIX** questions in all, selecting **ONE** question from each unit of Section B and all question in Sections A which is compulsory. Each question in Section B carries **12** marks and question in Section A carries **15** marks.

**SECTION—A**

1. (a) Distinguish between primary and secondary data. Give example of each.
- (b) Describe the limitations of graphic method of presenting statistical data.
- (c) Explain the meaning of correlation. State the extreme values of the coefficient of correlation and interpret them.
- (d) Explain the concept of regression. Why are there two regression lines in case of a bivariate data ?
- (e) What do you understand by (i) equally likely, (ii) mutually exclusive and (iii) independent events ?

- (f) What do you understand by conditional probability ?  
If  $P(A \cup B) = P(A) + P(B)$ . Are the two events A and B statistically independent ?
- (g) Define random variable and describe its probability distribution. Give an example.
- (h) Define probability mass function and probability density function. Give example of each.
- (i) State control limit theorem. Also write down any two of its uses.
- (j) Write down the uses of chi-square test.

## SECTION—B

### UNIT—I

2. (a) 'The standard deviation is most precise and the most satisfactory measure of dispersion.' Explain this statement by comparing the standard deviation with other measures of dispersion.
- (b) The mean and standard deviation of a sample of 100 observations were calculated as 40 and 5.1 respectively by a student who took by mistake 50 instead of 40 for one observation. Calculate the correct mean and standard deviation.
3. (a) What do you mean by quartiles ? Also define percentiles.

(b) Describe Box and Whisker diagrams.

(c) Draw a histogram for the following data :

Age (in years)	2—5	5—11	11—12
No. of boys	6	6	2
Age (in years)	12—14	14—15	15—16
No. of boys	5	1	3

### UNIT—II

4. (a) State and prove the multiplication theorem of probability. How is the result modified if the events are not independent ?

(b) What is the probability that at least two out of 100 people have the same birthday ? Assuming that all days of year are equally likely.

5. (a) The events  $E_1, E_2, \dots, E_n$  are mutually exclusive

and  $E = \bigcup_{i=1}^n E_i$ . Show that if  $P(A/E_i) = P(B/E_i)$ ;

$i = 1, 2, \dots, n$ , then  $P(A/E) = P(B/E)$ . Is this conclusion true if the events  $E_i$  are not mutually exclusive ?

(b) A restaurant serves two special dishes, A and B to its customers consisting of 60% men and 40% women, 80% of men order dish A and the rest B. 70% of women order dish B and rest A. In what ratio of A to B should the restaurant prepare the two dishes ?

## UNIT—III

6. (a) Prove that the regression lines of Y on X and X on Y intersect at the point  $(\bar{X}, \bar{Y})$ , Where  $\bar{X}$  and  $\bar{Y}$  are the means of X and Y.
- (b) From the following data between the ages of husbands and wives calculate the regression equations and find the husband's age when wife's age is 20 and wife's age when husband's age is 30.

7. (a) Explain clearly the scatter diagram method of measuring correlation. Also distinguish between positive and negative correlation with the help of it.

- (b) A computer while calculating the correlation coefficient between the variables X and Y obtained the following results :

$$N = 30, \Sigma X = 120, \Sigma X^2 = 600, \Sigma Y = 90, \Sigma Y^2 = 250, \Sigma XY = 356.$$

It was however, later discovered at the time of checking that it had copied down two pairs of observations as :

X	Y
8	10
12	7

while the correct values were

X	Y
8	12
10	8

Obtain the correct value of the correlation coefficient between X and Y.

## UNIT—IV

8. (a) Define mathematical expectation of random variables. Also describe the properties of expectation.
- (b) In a sequence of Bernoulli's trials, let  $X$  be the length of the run of either successes or failures starting with the first trial. Find  $E(X)$  and  $V(X)$ .
9. (a) Define the distribution function of a random variable and state its important properties.
- (b) A continuous random variable  $X$  has the distribution function :

$$F(x) = \begin{cases} 0 & , \text{ if } x \leq 1 \\ k(x-1)^4 & , \text{ if } 1 < x \leq 3 \\ 1 & , \text{ if } x > 3 \end{cases}$$

Find (i) value of  $k$ , (ii) probability density function  $f(x)$  and the median of  $X$ .

## UNIT—V

10. (a) What is difference between Bernoulli distribution and Binomial distribution ? Also obtain the recurrence relation for the central moments of Binomial distribution. Hence find its variance.
- (b) Show that in Poisson distribution with unit mean, mean deviation about mean is  $\frac{2}{e}$  times the standard deviation.

11. (a) Define normal distribution and find its first four central moments.

(b) To test the claim that the resistance of electric wire can be reduced by at least 0.05 ohm by alloying, 25 values obtained for each alloyed wire and standard wire produced the following results :

	Mean	Standard Deviation
Alloyed wire	0.083 ohm	0.003 ohm
Standard wire	0.136 ohm	0.002 ohm

Test at 5% level whether or not the claim is substantiated. ( $t_{48}(0.05) = -1.645$ )