Exam. Code : 206601
Subject Code : 4596

M.Sc. Bio-Informatics Semester—I

BASIC BIOSTATISTICS

Paper—BI–513

Time Allowed—3 Hours] [Maximum Marks—75

SECTION—A

Note :— Attempt **ALL** parts. Each part carries 1½ marks.

(10×1½=15)

1. (a) Discuss a frequency curve in brief.

(b) How will you calculate median in the case of ungrouped data ?

(c) Enunciate the demerits of classical definition of probability.

(d) Give the multiplication rule for the simultaneous occurrence of two events A and B.

(e) Give four examples where the use of regression analysis can be beneficially made.

(f) Interpret the values of correlation coefficient (ρ) as –1, 1 and 0.

(g) What is a probability mass function ?

(h) Define random variable and its mathematic expectation.

(i) Describe in brief the uniform distribution.

(j) When do you use paired t-test and how to apply it ?

SECTION—B

Note :— Attempt **ONE** question from each unit. Each question carries **12** marks.

UNIT—I

2. (a) Describe the different measures of central tendency of a frequency distribution.

   (b) For the following data, calculate :

   (i) Median

   (ii) Semi-inter quartile range and

   (iii) Standard deviation.

Wages in (Rs.)	No. of persons
170—180	52
180—190	68
190—200	85
200—210	92
210—220	100
220—230	95
230—240	70
240—250	28

3. (a) Describe shewness and kurtosis of the distribution. Also give their measures.

   (b) In a frequency distribution, the coefficient of skewness based upon the quartiles is 0.6. If the sum of the upper and lower quartiles is 100 and median is 38, find the value of the upper and lower quartiles.

UNIT—II

4. (a) Give concept and statistical (empirical) definition of probability. Also state and prove the law of total probability.

   (b) A player tosses a coin and is to score one point for every head and two points for every tail turned up. He is to play on until his score reaches or passes n. If p_n is the probability of attaining exactly 'n' score, show that $p_n = \frac{1}{2}\left(p_{n-1} + p_{n-2}\right)$ and hence find the value of p_n.

5. (a) Define conditional probability. If A and B are two events and $P(B) \neq 1$, prove that

$$P(A/\overline{B}) = \frac{P(A) - P(A \cap B)}{1 - P(B)}$$

where $\overline{B}$ denotes the complementary event to B and hence deduce that

$$P(A \cap B) \geq P(A) + P(B) - 1\,.$$

Also show that P(A) > or < P(A/B) according as $P(A/\overline{B}) > \text{or} < P(A)$.

(b) The probability that doctor A will diagnose a disease X correctly is 0.70. The prob. that a patient will die by his treatment after correct diagnosis is 0.35 and the prob. of death by wrong diagnosis is 0.75. A patient of doctor A who had disease X died. What is the probability that his disease was diagnosed correctly ?

UNIT—III

6. (a) Show that the coefficient of correlation is independent of change of scale and origin of the variable. How can you use scatter diagram to obtain an idea of nature of correlation coefficient.

(b) If U = ax + by and V = bx – ay, show that U and V are uncorrelated if

$$\frac{ab}{a^2 - b^2} = \frac{\rho \sigma_x \sigma_y}{\sigma_x^2 - \sigma_y^2}$$

where σ_x^2, σ_y^2 and ρ are variances of x, y and coefficient of correlation between them respectively.

7. (a) Define line of regression and regression coefficient. Also obtain the angle between the two lines of regression and interpret the cases when correlation coefficient (ρ) = 0 and $\rho = \pm 1$.

(b) The following results were obtained in the analysis of data on yield of dry bark in ounces (y) and age in years (x) of 200 cinchona plants :

	x	y
Average	9.2	16.5
Standard deviation	2.1	4.2

Correlation coefficient = + 0.84

Construct the two lines of regression and estimate the yield of dry bark of a plant of age 8.

UNIT—IV

8. (a) Define the distribution function of a random variable and state its important properties.

(b) A random variable x has the density function :

$$f(x) = k \cdot \frac{1}{1+x^2} \, y - \infty < x < \infty$$

Determine the value of k and distribution function. Evaluate $P(x \geq 0)$.

9. Show that mathematical expectation of the sum of two random variables is the sum of their individual expectations. If two variables are independent, the mathematical expectation of their product is the product of their expectations.

UNIT—V

10. (a) Define binomial distribution and find its mean, variance and skewness.

(b) The probability of a man hitting a target is $\frac{1}{3}$. How many times must he fire so that the probability of hitting the target at least once is more than 90% ?

11. (a) Define normal distribution. Also show that normal distribution is the limiting form of binomial distribution.

(b) A survey of 800 families with four children each revealed the following distribution :

No. of boys :	0	1	2	3	4
No. of girls :	4	3	2	1	0
No. of families :	32	178	290	236	64

To this result consistent with the hypothesis that male and female birth are equally probable ?

$(\chi^2_{4;0.05} = 9.488)$.