Exam.Code:5142
Sub. Code: 14313

2055

B.Sc. (Hons.) Bio-Informatics (FYUP)

Second Semester

BINF -2003: Statistical Methods in Bioinformatics

Time allowed: 3 Hours

Max. Marks: 60

NOTE: Attempt five questions in all, including Question No. 1 which is compulsory and selecting two questions from each Unit. Candidates may ask for Statistical tables from the Superintendent of Examination Centre. Simple Non-Programmable calculator is allowed.

x-x-x

1. (a) Differentiate between classification and tabulation of data.
   (b) Explain Box and Whisker plot.
   (c) Write the ANOVA table for one way classification.
   (d) Explain type-1 and type-2 errors in testing.
   (e) Define conditional probability for two events.
   (f) Define the level of significance and critical region (6×2)

Unit-I

2. (a) Discuss the various methods of collecting primary and secondary data.

   (b) Draw histogram and frequency polygon of the following data:

Class	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	12	16	19	17	14	12

(6, 6)

3(a) Find the first four moments about mean (central moments) of the following data.

Class	5	15	25	35	45	55
Frequency	2	6	9	7	4	2

(b) Find the median, Mode and standard deviation for the following data:

X:	0-10	10-20	20-30	30-40	40-50	50-60
F:	5	7	11	16	9	4

(c) What is skewness? Draw rough sketches to indicate different types of skewness. (4, 6, 2)

4. (a) Suppose 70 per cent of all tourists who come to India will visit Agra while 60 per cent will visit Goa and 50 per cent of them will visit both Agra and Goa. What is the probability that a tourist will visit either Goa or Agra or both?

(b) Differentiate between probability mass function and probability density function.

Contd... P/2

(c) A random variable X has the following probability function:

Value of X, x:	0	1	2	3	4	5	6	7	8
$p(x)$:	K	3k	5k	7k	9k	11k	13k	15k	17k

(i) Find the constant k, (ii) Evaluate $P(X<5)$, $P(X\geq 5)$, *and* $P(0<X<6)$, and (iii) Determine the distribution function F(x) of X. (4, 4, 4)

Unit-II

5(a) Define the followings:
(i) Test of independence of attributes.
(ii) Chi-square test of goodness of fit.

(b) It is claimed that a random sample of 100 tyres with a mean life of 15269kms is drawn from a population of tyres which has a mean life of 15200kms and a standard deviation of 1248kms. Test the validity of the claim at 5% level of significance. (8, 4)

6(a) Develop a test procedure for testing the equality of two populations proportions, when the two populations distributed normally.

(b) Below are given the gain in weights (in kgs.) of pigs fed on two diets A and B.

Gain in weight

Diet A : 25, 32, 30, 34, 24. 14. 32, 24, 30. 31, 35, 25
Diet-B : 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22
Test. if the two diets differ significantly as regards their effect on increase in weight (Given $\alpha = 5\%$). (6, 6)

7(a) A manufacturing company has purchased three new machines of different makes and wishes to determine whether one of them is faster than the others in producing a certain product. Five hourly production figures are observed at random from each machine and the results are given in table below:

Machine A	25	30	36	38	31
Machine B	31	39	38	42	35
Machine C	24	30	28	25	28

Use analysis of variance and determine whether the machines are significantly different in their mean speeds. (Use $\alpha = 5\%$).

(b) Differentiate between null and alternative hypothesis with suitable examples. (8, 4)

x-x-x