Goa Vidyaprasarak Mandal's
Gopal Govind Poy Raiturcar College of Commerce and Economics
Ponda - Goa
B.C.A. (Semester - IV) Examination, April 2017

404 DATA ANALYSIS AND STATISTICAL TECHNIQUES

Duration : 2 Hrs Marks : 50

Instructions:

I. All the questions are compulsory however internal choices are given.

II. Use of non programmable calculators is allowed.

III. Marks to the right indicate full marks.

Q.1. I) Fill in the blanks with the correct alternatives given in the bracket: (5)

i) ..... is a small group of individuals selected from the population such that it possesses almost all the characteristics of the population.
(sample, attribute, survey)

ii) ..... is a quantitative characteristic of an individual of a population.
(attribute, variate, frequency)

iii) The coefficient of correlation 'r' indicates a positive correlation between x and y if $r \approx$ ..... (0, 1, 2)

iv) The regression coefficient of 'y on x' is given by .....
(b_{yx}, cov(x,y), $\boldsymbol{\sigma}_x\boldsymbol{\sigma}_y$)

v) Two events with sample space A and B resp. are such that $A \cap B = \phi$.
The two events are .....(mutually exclusive, exhaustive, complementary)

Q.1.II) Answer the following: (5)

i) A die is thrown. What is the probability of getting a prime number?

ii) Give an example of a continuous random variable.

iii) What is data mining?

iv) Write the probability density function for Poisson distribution with parameter λ.

v) We want to verify whether a coin is unbiased. Set up the null hypothesis for the same.

Q.2.

i) Find the quartiles $\mathbf{Q_1}$ and $\mathbf{Q_2}$ for the following data : (5)

Weight (in kg)	45-50	50-55	55-60	60-65	65-70	70-75
No. of students	16	17	20	21	14	12

ii) Draw histogram for the following : (3)

C.I.	0-100	100-200	200-300	300-400	400-500	500-600	600-700	700-800
F	50	30	70	20	10	40	50	30

iii) Find the weighted mean for the following : (2)

x	20	15	10	14	12
w	6	4	3	3	4

OR

Q.II.

i) Calculate the variance and the standard deviation for the following (5)

Class intervals	0-2	2-4	4-6	6-8	8-10
Frequency	10	20	30	10	10

ii) Find the percentiles $\mathbf{P_{15}}$ and $\mathbf{P_{25}}$ for the following (5)

Age(in years)	15-20	20-25	25-30	30-35	35-40	40-45
No. of persons	6	17	15	25	5	7

Q.3.

i) The average number of incoming telephone calls at a switch board per minute is 2. Find the probability that during a given minute, two or more calls are received. (e^{-2} =0.135) (5)

iii) For the following data, obtain the equations of regression line of 'x on y' and hence determine the most likely value of x when y=4.5 (5)

x	2	3	4
y	4	5	2

OR

Q.III.

i) The ranking of 8 individuals at the start and at the finish of a course of training are as follows :

Individual	A	B	C	D	E	F	G	H
Rank before	1	2	3	4	8	5	4	2
Rank after	2	3	1	4	3	6	7	6

Calculate Spearman's coefficient of correlation. (5)

ii) If a random variable X follows Poisson distribution such that P(1)=P(2). Find its mean and variance. (5)

Q.4.

i) Two unbiased dice are rolled. Find the probability that the sum of the numbers on the two faces is either divisible by 2 or 3. (5)

ii) What are the steps in the knowledge discovery process. (5)

OR

Q.IV.

i) A card is drawn at random from a well shuffled pack of 52 cards. Find the probability that it is

(*a*) a red card or a king, (*b*) a black ace, (*c*) an ace and a jack (5)

ii) Give applications of data mining. (5)

Q.5.

i) An automatic can filling machine on an average, fills 180 ml of milk with a standard of 2 ml. Find the probability that the average volume of milk filled in 100 cans from a lot is at most 180.2 ml. (Given Z=0 to Z=1 is 0.3414) (5)

ii) An oceanographer finds from the past records that the average depth of an ocean in a certain region is 56.9 fathoms with a standard deviation of 4.6 fathoms. He decides to check the value of mean depth by selecting a sample of 34 (measuring the depth at 34 different points) at random and then testing the validity of the mean at 1% level of significance. If he finds the sample mean of 59.3 fathoms, what would be his conclusion? (5)

OR

Q.V.

i) A random sample of size 400 has sample proportion 0.75. Can we say that, it is drawn from a population with a proportion P=0.8 at 5% level of significance? (5)

ii) The weekly wages of 1000 workers are normally distributed with mean `900 and standard deviation ` 50. Estimate the number of workers whose weekly wages will be between ` 900 and ` 1000.
(Given Z=0 to Z=2 is 0.4772) (5)

xxxxxxxxxxxxxxxxx