[16/A-11]

SARDAR PATEL UNIVERSITY
B.Sc. Semester – IV Examination
Friday, 20th April, 2018

Time: 10.00 to 12.00 p.m   US04FSTA01   M.Marks: 70

(Foundation of Statistics – II)

Note: (i) Statistical table will be allowed /provided on request (ii) Simple/Scientific calculator is allowed.
(iii) Q.3 to 6 each sub question is of 5 marks.

Q.1 Multiple Choice Questions (10 × 1)

(1) If $\frac{6 \sum di^2}{n(n^2 - 1)}$ is zero then the rank correlation coefficient be

    (a) 0.5   (b) −1   (c) 1   (d) Impossible to calculate

(2) Use the standard normal distribution to find $P(-2.25 < Z < 1.25)$

    (a) 0.8822   (b) 0.4878   (c) 0.8944   (d) 0.8786

(3) When testing for independence in a contingency table with 3 rows and 4 columns, there are _____degrees of freedom.

    (a) 2   (b) 4   (c) 6   (d) 8

(4) Which of the following probability distribution has mean and variance always same?

    (a) Binomial   (b) Poisson   (c) Normal   (d) None of the above

(5) The ranks given by two different judges to five participants in a debate contest are:

R1	5	4	3	2	1
R2	1	2	3	4	5

The rank correlation coefficient between them is

    (a) 0   (b) 1   (c) −1   (d) Can't possible

(6) Consider the following probability distribution:

$$P(X = x) = \frac{e^{-2} 2^x}{x!}, x = 0, 1, 2 \ldots$$

The value of $P(X < 3) =$ _____

    (a) 0.8571   (b) 0.1429   (c) 0.3233   (d) 0.6767

(7) In a standard normal distribution, the area to the left of $Z = 1$ is

    (a) 0.6413   (b) 0.7413   (c) 0.8413   (d) 0.9413

(8) Which of the following is true?

    (a) $r = \pm\sqrt{b_{XY} \times b_{YX}}$ (b) $-1 \le r \le 1$ (c) two regression lines intersect at $(\bar{X}, \bar{Y})$ (d) All of the above

(9) Coefficient of skewness for Normal distribution is

    (a) > 0   (b) < 0   (c) 0   (d) Undefined

(10) In the regression equation $Y = 3X + 4$, what does the 4 represent?

    (a) Slope of the line   (b) Y intercept

    (c) Any value of the independent variable that is selected   (d) None of the above

Q.2 Short Type Questions (Attempt Any Ten) (10 × 2)

(1) Two regression equations are $Y = -3.75 + 1.25X$ and $X = -1.47 + 0.36Y$. Find the coefficient of correlation between X and Y.

(2) The length of the pregnancy duration of human being is assumed to be normally distributed with a mean of 266 days and a standard deviation of 16 days. If a single birth was randomly selected, what is the probability that her pregnancy duration is more than 280 days?

(3) State the nature of the following correlations:

(i) Age of applicant for life insurance and the premium of insurance

(ii) The colour of sari and the intelligence of the lady who wears it.

(4) Following table gives the number of those plants which have some special qualities. To test the hypothesis that the colour of flowers is independent to the flatness of leaves. Which statistical test would you recommend? Write in brief about it.

	Type of leaves		Total
	Flat leaves	Curved leaves	
White flowers	90	36	126
Red flowers	29	5	34
Total	119	41	160

(5) Write down the properties of Normal distribution.

(6) How will you calculate rank correlation coefficient? If ranks are repeated how will you modify it?

(7) Find the area of the indicated region under the standard normal curve.
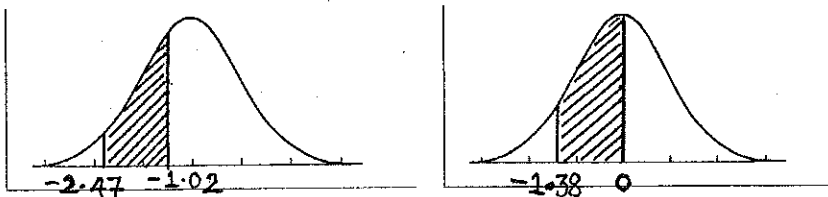
(8) The mean and standard deviation of binomial distribution are 9 and 2 respectively. Find $P(X \geq 2)$.

(9) Define Binomial distribution. State the conditions under which Binomial distribution tends to Poisson distribution.

(10) Write down the regression equation of X on Y and Y on X. State its uses.

(11) What is regression? Write down the properties of regression coefficients.

(12) The probability that a patient will get reaction of a temiflu injection is 0.02. If 110 patients are given that injection, find the probability that 3 or more will get reaction from that injection.

Q.3(a) Do as directed:
(i) Define correlation coefficient. Interpret the cases when $r = -1, 0, 1$
(ii) Write a note on scatter plot (diagram) with its limitations, if any.

(b) An economist wanted to analyze the relationship between the speed of a car and its mileage. An experiment was carried out at different speeds (KMPH) and mileage (KMPL) was recorded.

Speed (X)	25	35	45	50	60	65	70
Mileage (Y)	40	39	37	33	30	27	25

* KMPH – Kilometer per hour and KMPL – Kilometer per liter
(i) Identify the objective (s) of the experiment and accordingly select an appropriate statistical measure and calculate it. (ii) Find the regression equation which could be used to predict the mileage of a car when the speed is 55 KMPH?

OR

Q.3(a) List out the various methods of studying relationship between two variables. According to you, which methods do you considered to be best and why?

(b) The following table consists of one student athlete's time (in minutes) to swim 2000 yards and the student's heart rate (beats per minute) after swimming on a random sample of 10 days.

Swim time	34.12	35.72	34.72	34.05	34.13	35.73	36.17	35.57	35.37	35.57
Heart rate	144	152	124	140	152	146	128	136	144	148

(*i*) Identify an independent and dependent variable (*ii*) Does there appear to be any evidence of linear relationship between these two variables? Justify your answer by calculating most suitable statistical measure. (*iii*) Estimate the heart rate of a student if his swim time is 35.42 minutes.

Q.4(a) Let X be a Poisson variate with variance 2. Determine the following probabilities:

(*i*) $P(X < 2)$ (*ii*) $P(X = 3)$ (*iii*) $P(X > 4)$ (*iv*) $P(X \geq 1)$

(b) A medication gives 10% of the users an undesirable reaction. If a sample of 15 users receives the medication, find the prob. of (*i*) at least 2 (*ii*) exactly 4 (*iii*) more than 3, will have undesirable reactions.

OR

Q.4(a) It is known that 15% of plants produced by a certain species of corn seed will be infertile. In a random sample of 10 such plants, what is the probability that (*i*) exactly 2 (*ii*) more than 3, will be infertile?

(b) A viral infection is spread by contact with an infected person. Let the probability that a healthy person gets an infection, in one contact, be 0.05.

(*i*) An infected person has contact with 150 healthy persons. Specify the distribution of X,

Where X = no. of healthy persons who contract an infection (*ii*) Find $P(X \leq 3)$, $P(X = 0)$.

Q.5(a) Given that Z is a standard normal variable, Sketch each one and evaluate the following probabilities.

(*i*)$P(Z \geq -2.02)$ (*ii*) $P(Z < 1.86)$ (*iii*) $P(Z \geq -1.02)$ (*iv*) $P(-1.62 \leq Z \leq 1.62)$

(b) The mean and standard deviation of marks of 500 students in an examination are 52 and 8 respectively. If the distribution of the marks is approximately normal, Find % of Students getting marks (*i*) more than 60 (*ii*) between 48 and 56 (*iii*) less than 40.

OR

Q.5(a) The measurement of the length of the index finger of a human right hand is a normally distributed with a mean of 6 cm. and a standard deviation of 0.5 cm. What is the probability that the finger length of a randomly selected person will be (*i*) between 4.8 cm. and 7.2 cm (*ii*) more than 4.7 cm (*iii*) less than 5.6 cm.

(b) A machine produces components, the lengths of which are normally distributed with mean 102.30 mm and s.d 1.20 mm. What is the prob. that a randomly selected component has length (*i*) Less than 102 mm (*ii*) more than 104.7 mm?

Q.6 A sample of 300 students of Under Graduate (UG) and 300 students of Post Graduate (PG) of a university were asked to give their opinion towards the autonomous status of colleges. 190 of the Under Graduate and 210 of the Post Graduate students favored the autonomous status. Present the above data in a tabular form and test at $\propto= 0.05$ that opinions of Under Graduate and Post Graduate students on autonomous status of colleges are independent.

OR

Q.6 A manager of a sports club collects information regarding the sports in which members played and their respective ages. 643 members of the sports club were randomly selected and following information was received:

Sports	Age Groups				Total
	18 – 25	26 – 30	31 – 40	41 and over	
Badminton	42	58	30	46	176
Table Tennis	58	76	38	65	237
Swimming	72	60	65	33	230
Total	172	194	133	144	643

Test whether there is a relationship between the age of a member and his/her choice of sport. Test at $\alpha = 0.05$.

————— X —————