

[17/A-10]

SEAT No. \_\_\_\_\_

No. of Printed Pages : 03

SARDAR PATEL UNIVERSITY  
B.Sc. Semester – IV Examination  
Friday, 12<sup>th</sup> April, 2019  
US04ESTA04  
(Biostatistics – II)

Time: 10.00 to 12.00 p.m

M.Marks: 70

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) For a random sample of 9 women, the average resting pulse rate is  $\bar{X} = 76$  beats per minute, and the sample standard deviation is  $s = 5$ . the standard error of the sample mean is \_\_\_\_\_  
(a) 0.557 (b) 0.745 (c) 2.778 (d) 1.667
- (2) The degrees of freedom for a t – test of a mean if the sample size is 10 is  
(a) 8 (b) 9 (c) 10 (d) 11
- (3) The area under the normal curve between  $z = -2$  and  $z = 0$  is \_\_\_\_\_ the area under the normal curve between  $z = 0$  and  $z = 1$ .  
(a) Less than (b) Greater than (c) Equal to (d) None of these
- (4) A group of 10 men were given a special diet for two weeks to test weight loss in pounds. The observed data was:

Man	1	2	3	4	5	6	7	8	9	10
Weight before	181	171	190	187	210	202	166	173	183	184
Weight after	178	172	185	184	201	201	160	168	180	179

To determine if the data provide sufficient evidence to indicate the special diet leads to a weight loss, the appropriate test procedure is :

- (a) Paired t - test (b) Unpaired t – test (c) Z - test (d) All of the above
- (5) The choice of one – tailed and two – tailed test depends on  
(a) Null hypothesis (b) Alternative hypothesis (c) Composite hypothesis (d) None of these
- (6) In the regression equation  $Y = 21 - 17X$ , 21 is the \_\_\_\_\_ while  $-17$  is the \_\_\_\_\_ of the regression line  
(a) radius, intercept (b) intercept, slope (c) slope, intercept (d) slope, regression coefficient
- (7) When testing for independence in a contingency table with 3 rows and 5 columns, there are \_\_\_\_\_ degrees of freedom.  
(a) 8 (b) 4 (c) 6 (d) 12
- (8) For a normal distribution  
(a) Mean = Median = Mode (b) Coefficient of skewness is zero  
(c)  $Q_2 = \frac{Q_1+Q_3}{2}$  (d) All of the above
- (9) If all the points in a scatter diagram lie on the least squares regression line, then the coefficient of correlation must be  
(a) 1 (b)  $-1$  (c) 0 (d) either  $-1$  or 1
- (10) 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

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) If  $X \sim N(0, 1)$ , find the value of  $k$  such that  $P(X \leq k) = 0.7019$ .
- (3) 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

(1)

- (4) What is regression? State its uses.
- (5) In testing hypotheses  $H_0 : \mu = 75$  v/s  $H_1 : \mu \neq 75$ , the following information is known:  $n = 12$ ,  $\bar{X} = 78$  and  $S = 10$ . Carry out an appropriate statistical test to test the hypotheses and comment on it.
- (6) Let  $X \sim N(\mu, \sigma^2)$  such that  $Q_1 = 12$  and  $Q_3 = 22$  then find  $\mu$ .
- (7) Give two examples each of (i) positive correlation (ii) Negative correlation, applicable in the field of Biosciences.
- (8) Write in brief on chi square test in a  $2 \times 2$  contingency table.
- (9) 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.
- (10) Two diets were to be compared. Seventy five individuals were selected at random from a population of overweight people. Forty of this group were assigned to diet A and the remaining thirty five were placed on diet B. The weight losses in pounds over a period of one week were found and the following information was recorded.

	Sample size	Sample mean(lbs)	Sample variance
Diet - A	40	10.3	7.00
Diet - B	35	7.3	3.25

Identify an objective(s) of the study. Which statistical test would you recommend? Write in brief about it.

- (11) Define two types of errors in the testing of hypothesis.
- (12) 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?

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.  
 (iii) Find the regression equation which could be used to predict the mileage of a car when the speed is 55 KMPH?  
 (iv) List out the variables which may influence the mileage of a car.

OR

Q.3(a) Write down the properties of correlation coefficient and regression coefficients.

- (b) A biologist assumes that there is a linear relationship between the amount of fertilizer supplied to tomato plants and the subsequent yield of tomatoes obtained. Eight tomato plants, of the same variety, were selected at random and treated, weekly with a solution in which X grams of fertilizer was dissolved in a fixed quantity of water. The yield, Y kilograms, of tomatoes was recorded.

Plant	1	2	3	4	5	6	7	8
X	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5
Y	3.9	4.4	5.8	6.6	7.0	7.1	7.3	7.7

- (i) Plot a scatter diagram and comment on the type of relationship that exists between these two variables.  
 (ii) Calculate the equation of the least squares regression line of Y on X. (iii) Estimate the yield of a plant treated, weekly, with 3.2 grams of fertilizer.

Q.4(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. If 1000 humans are randomly selected, how many of them will have right hand finger length (i) between 4.8 cm. and 7.2 cm (ii) more than 4.7 cm (iii) less than 4.7 cm.

- (b) A new tax law is expected to benefit "middle income" families, those with incomes between \$20,000 to \$30,000. If family income follows normal distribution with mean \$25,000 and standard deviation \$10,000, what percentage of the population will benefit from the new tax law?

OR

- Q.4(a) Total fingerprint ridge count in humans is approximately normally distributed with mean of 140 and s.d of 50. Find the prob. that an individual picked at random will have a ridge count (i) Less than 100 (ii) between 100 to 145 (iii) more than 132.
- (b) The lengths of certain species of fish are normally distributed with mean 60 cm. If 80% of fish are of length more than 42 cm; determine the value of standard deviation.
- Q.5 A survey of 200 workers was conducted regarding their education and the level of their job satisfaction. The results are as under:

Level of education	Level of job satisfaction		
	Low	Medium	High
School	20	35	25
College	17	33	20
Post Graduate	11	18	21

Test at 5% level of significance whether the level of job satisfaction depends on the level of education.

OR

- Q.5 To determine the possible effect of a chemical treatment on the rate of seed germination, 100 chemically treated seeds and 150 untreated seeds are sown. The number of seeds that germinated is recorded. Is the treatment effective? Test at  $\alpha = 0.05$ .

Types of seed	Germinated	Not germinated	Total
Treated	84	16	100
Untreated	132	18	150
Total	216	34	250

- Q.6(a) To compare the effectiveness of Isometric and Isotonic exercise methods in abdominal reduction, 20 potbellied business executives are included in an experiment. 10 are selected at random and assigned to one exercise method; the remaining 10 are assigned to the other exercise method. After 5 weeks, the reductions in abdomen measurements are recorded in centimeters, and the following results obtained.

	Isometric Method	Isotonic Method
Mean	2.5	3.1
Standard Deviation	0.8	1.0

Do these data support a claim that Isotonic method is more effective? Test at  $\alpha = 0.05$ .

- (b) A city health department wishes to determine if the mean bacteria count per unit volume of water at a lake beach is within the safety level of 200. A researcher collected samples of unit volume and found the bacteria count to be

175	190	215	198	184
207	210	193	196	180

Does the data indicate that the bacteria count is within the safety level? Test at  $\alpha = 0.05$ .

OR

- Q.6 A gym instructor wanted to know the impact of a fitness schedule on the fat percentage of male clients. A study was planned in which 12 men were randomly selected. Their fat percentage was measured before they joined the gym. After 12 - week schedule, they were again tested for their fat percentages. The data obtained shown below:

Before	26	18	29	18	16	18	19	18	21	22	24	19
After	22	17	26	19	17	17	16	17	20	22	21	20

On the basis this data, can it be concluded that the fat percentage of the clients has significantly reduced as a result of the fitness schedule? Test at  $\alpha = 0.05$ .

— \* —  
(3)

