[18&A-19]

Sardar Patel University

B.Sc. Semester - IV Examination

Tuesday, 10th April, 2018

Time: 10 .00 to 1.00 pm   Course Code: USC04CSTA01   M.Marks: 70

(Statistical Techniques)

Note: (i) Simple/Scientific calculator is allowed.   (ii) Graph paper will be provided on request.

(iii) Figures to the right indicate marks.   (iv) Q.3 to 6 each sub question is of 5 marks

----

Q.1 Multiple Choice Questions   (10×1)

(1) In a linear regression, the least squares line is $\hat{Y} = -3.75 + 1.25X$, and the coefficient of determination is 0.81 then the coefficient of correlation must be

(a) 0.90   (b) −0.90   (c) 1.25   (d) −3.75

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

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

The rank correlation coefficient between them is

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

(3) Which type of control chart should be used when it is possible to have more than one mistake per item

(a) p − chart   (b) np − chart   (c) $\bar{X}$ − chart   (d) C − chart

(4) The points falling outside the control limits shows presence of_____causes of variations

(a) Assignable   (b) Chance   (c) Random   (d) None of these

(5) In semi − average method, if the time series data contains odd number of values then we drop

(a) First Value   (b) Last Value

(c) Middle Two Values   (d) Middle Value

(6) 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

(7) With reference to SQC, the seven consecutive point lie on either side of central line is said to be

(a) Run   (b) Run above   (c) Run below   (d) $\frac{1}{2^6}$

(8) Twenty samples of size 5 are taken from a production process. The mean of all sample means is 42.5 and the mean range of the samples is 1.5. What is the UCL for the $\bar{X}$ − chart?

(a) 3.17   (b) 1.5   (c) 0   (d) 43.37

(9) In linear regression, which of the following statement indicate that there is no linear relationship between two variables X and Y

(a) Coefficient of determination is 1   (b) Coefficient of correlation is zero

(c) Sum of squares of errors is zero   (d) None of the above

(10) The following table contains the number of complaints received in a Government department for the first 6 months of a year.

Month	Jan	Feb	Mar	Apr	May	Jun
No. of complaints	36	45	81	90	108	144

If a 3 − month moving average is used to calculate trend, what would be the trend value for the month of March?

(a) 36   (b) 40.5   (c) 54   (d) 72

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

(1) The regression lines of Y on X and X on Y are $Y = a + bX$ and $X = c + dY$. find the correlation coefficient between X and Y.

(2) Determine $'a'$ and $'b'$ in the equation $Y = a + bX$ to the following data.

X	2	4	6	8	10
Y	3.07	12.85	31.47	57.38	91.29

(3) Explain the meaning of Low spot points with reference to C − chart?

(P.T.O.)

(4) The least squares trend line for an annual time series data regarding sale of cars (In 000 units) from 2010 to 2017 is $Y = 1.2 + 0.5X$. Identify the intercept and slope of this trend line. Interpret the slope in this trend line.

(5) What are the chief sources of assignable causes of variations?

(6) Find the correlation coefficient between X and $a - X$, where 'a' is any positive constant.

(7) Why do we analyze Time Series?

(8) With reference to SQC explain in brief about Run above and Run below.

(9) Write down the normal equations for estimating unknown constants in the equation $Y = a + bX + cX^2$.

(10) If one or more points fall below LCL in construction of $p -$ chart, what would you conclude from that?

(11) What is Time series data? Give some examples you may have encountered in everyday life.

(12) Write down the properties of regression coefficients. Prove any one of them.

Q.3 (a) Explain the meaning of coefficient of correlation and rank correlation coefficient. If the coefficient of rank correlation is 1, is it necessary that the coefficient of correlation is also 1? Give the counter example in support of your answer.

(b) Fit an equation $Y = a + bX$ to the data given below and estimate sex - ratio for the year 2011.

Sex – ratio of Gujarat from 1901 to 2001

Year	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991	2001
Sex – ratio	954	946	944	945	941	952	940	934	942	934	921

Source: www. planningcommission.nic.in

OR

Q.3 (a) Using least squares method obtain normal equations for estimating unknown constants in the equation $Y = a + bX + cX^2$.

(b) In usual notation, Prove that r lies between -1 and $+1$. Interpret the cases when $r = -1, 0, 1$.

Q.4 (a) Do as directed:

(i) Prove that if one of the regression coefficient is greater than one than the other one must be less than one.

(ii) The tangent of an angle between two regression line is given to be 0.6 and $S_Y = 2S_X$ then find the correlation coefficient between X and Y.

(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.

OR

Q.4 (a) Do as directed:

(i) Prove that the arithmetic mean of two regression coefficients is always greater than the correlation coefficient.

(ii) Why there are two regression lines? Under which condition(s) two regression lines interchangeable?

(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.

Q.5 (a) What is time series? State its importance. Why do we analyze time series? Give its component and discuss each of them.

(b) Calculate the quarterly trend values by trend method for the following data:

Quarter	Sale of Cars (in 000 units)				
	2013	2014	2015	2016	2017
I	45	48	49	52	60
II	54	56	63	65	70
III	72	63	70	75	83
IV	60	56	65	72	86

OR

Q.5 (a) Which components of a time-series would you mainly associate each of the following? Why?

(*i*) A strike in a factory delaying production of a car in a Manesar plant of Maruti Udyog.

(*ii*) Price hike in Gold

(*iii*) Fall in death rate due to advances in Science

(*iv*) The sale of air conditioners increases during summer.

(b) From the following data, find the trend values by the method of semi - averages. Also, estimate the Facebook users for 2019. Also find average monthly increase in the users.

Number of Facebook users in India from 2011 to 2018 (in Millions)

Year	2011	2012	2013	2014	2015	2016	2017	2018
Facebook Users	135.6	165.57	194.11	219.94	242.53	261.83	277.95	291.5

Q.6 (a) State the various causes of variations in the production process. Write in brief about them.

(b) Do as directed:

(*i*) Write down the uses of SQC (*ii*) Theory of Run

OR

Q.6 (a) Differentiate between (*i*) p and np chart (*ii*) Variable and Attribute chart.

(b) A company manufactures valves for industrial use. Ten samples of 150 valves each were taken from the production line and tested. The results are reported below:

Sample No.	1	2	3	4	5	6	7	8	9	10
No. of defectives	32	18	07	16	09	29	32	13	18	07

(*i*) Construct control chart for no. of defectives using 3σ limits. (*ii*) Is the process in control?

———X———