Sardar Patel University
B.Sc. Semester – IV Examination (NC 2010 Batch)

Time: 10.00 to 12.00 am   Friday, 5th April, 2019   Maximum Marks: 70
Foundation of Statistics – II
(US04FSTA01)

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

Q.1   Multiple Choice Questions (1×10)
(1)   The ranks according to two attributes say I.Q. and Beauty are given below:

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

Where R1 and R2 are ranks assigned to the participants based on their I.Q and Beauty. The rank correlation coefficient be
(a) 1   (b) −1   (c) 0   (d) Not possible

(2)   Which of the following probability distribution always has mean and variance be equal?
(a) Binomial   (b) Poisson   (c) Normal   (d) All of the above

(3)   If $\frac{6 \sum d_i^2}{n(n^2 - 1)}$ is zero then the rank correlation coefficient be
(a) 0.5   (b) −1   (c) 1   (d) Impossible to calculate

(4)   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

(5)   Let $X \sim P(3)$ then $P(X > 2) = $ _____
(a) 0.4232   (b) 0.1992   (c) 0.6472   (d) 0.5768

(6)   Which of the following is not true?
(a) $-1 \le r \le 1$ (b) $b_{XY} = r \frac{s_x}{s_y}$ (c) $r = \pm \sqrt{b_{XY} \times b_{YX}}$ (d) two regression lines do not intersect

(7)   Consider the following probability distribution:
$P(X = x) = \binom{10}{x}(0.3)^x (0.7)^{10-x}, x = 0, 1, \dots, 10$ then $P(X \le 10) = $ _____
(a) 0   (b) 0.21   (c) 1   (d) 0.45

(8)   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

(9)   The regression equation of Y on X is $Y = -1.75X + 0.45$ then
(a) $b_{XY} = -1.75$ (b) $b_{XY} = 0.45$ (c) $b_{YX} = -1.75$ (d) $b_{YX} = 0.45$

(10) 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

Q.2   Short Type Questions (Attempt Any Ten) (2×10)
(a)   The number of customers entering a bank per minute is a Poisson variate with a mean of 4 customers per minute. What is the probability that less than two customers enter the bank in minute?
(b)   The vitamin content of a particular brand of vitamin supplement pill is normally distributed with mean 490 mg and s.d 64 mg. What is the probability that a randomly selected pill contains atleast 420 mg of vitamin?
(c)   Define Poisson distribution. Give some examples where Poisson distribution may be used.
(d)   What is scatter plot (diagram)? Write down its limitations, if any.
(e)   Write in brief on chi square test in a 2×2 contingency table.
(f)   Write down the properties of Normal distribution.
(g)   What is regression? State its uses.
(h)   Two regression equations are $Y = -3.75 + 1.25X$ and $X = -1.47 + 0.36Y$. Find the coefficient of correlation between X and Y.
(i)   The probability that a patient will get reaction of a temiflu injection is 0.20. If 15 patients are given that injection, find the probability that 2 or less will get reaction from that injection.
(j)   Give two examples each of

(i) Positive correlation (ii) Negative correlation (iii) Spurious (Non - sense) correlation.

(k) Define Binomial distribution. State the conditions for applicability of Binomial distribution.

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

Q.3(a) What is correlation? List out the various methods of studying correlation. Write in brief about any one of them.

(b) A random sample of seven drivers insured with a company and having similar auto insurance policies was selected. The following table lists their driver experiences (in years) and monthly auto insurance premiums (in dollars)

Driving Experience	5	2	12	9	15	6	25
Monthly Insurance Premiums	64	87	50	71	44	56	42

(i) Does the insurance premium depend on driving experience? Justify your answer by calculating most suitable statistical measure (ii) Predict the monthly auto insurance premium for driver with 14 years of driving experience.

OR

Q.3(a) What is regression? Write down the properties of regression coefficients? Write down the regression equation which could be used to predict Y for any given values of X.

(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 and 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) A life insurance agent sells on the average 3 life insurance policies per week. Use Poisson distribution to calculate the probability that in a given week he sells (i) some policies (ii) 2 or more policies

(b) An institute found that 10% of the registered students withdraw without completing a course on C++. If 12 students have registered in the current batch, compute the probability that (i) More than 3 (ii) Exactly 4, will withdraw.

OR

Q.4(a) It was claimed that 20% dentists recommend Colgate sensitive tooth paste to his patients in sensitivity of teeth. Suppose that the claim is true. If 15 dentists are selected independently and at random. Let X be the no. of dentists who recommend Colgate sensitive paste to his/her patients. Name the distribution of X and state its mean and standard deviation. Calculate (i) $P(X > 3)$ (ii) $P(X < 2)$

(b) The probability that a patient will get reaction of a temiflu injection is 0.01. If 150 patients are given that injection, find the probabilities that (i) Exactly 3 (ii) less than 2 (iii) at least 4 (iv) none, will get reaction.

Q.5(a) Pulse rates of adult men are normally distributed with a mean of 70 and standard deviation of 8. Find the probability that a randomly selected man have pulse rate (i) Greater than 77 (ii) Between 62 to 72 (iii) Less than mean.

(b) Given that Z is a standard normal variable, Sketch each one and evaluate the following probabilities.
$(i)P(Z \geq 1.54)(ii)P(Z \leq -1.08)(iii)P(Z \geq -1.27)(iv)P(-1.26 \leq Z \leq 1.62)(v)P(2.03 \leq Z \leq 0)$

OR

Q.5(a) 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 normal, Find % of Students getting marks (i) more than 60 (ii) between 48 and 56 (iii) more than mean

(b) Let X be a normal variate with mean 5 and standard deviation 3. Determine the following probabilities:
(i) $P(X > 3.2)$ (ii) $P(X < 2.5)$ (iii) $P(-1.2 < X < 3.07)$ (iv) $P(2.7 < X < 5.3)$ (v)$P(X \leq -2.3)$

Q.6 A plant breeder wants to know if the sterility of rice is a genetic problem. Samples were taken from a large field study of 400 plots and the sterility of each plot was rated as follows:

Sterility	Genotypes			
	A	B	C	D
No problem	20	15	12	10
Moderate	70	60	80	50
severe	10	25	8	40

Test the hypothesis that the severity of sterility is independent of genetic make – up or genotypes.

OR

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 $\alpha = 0.05$ that opinions of Under Graduate and Post Graduate students on autonomous status of colleges are independent.