State all your assumptions very clearly. This will carry weightage.
There are five questions. Attempt *any FOUR* questions.
Marks: 80   Weightage : 40 %   Time: 2 Hours

Q1 (a) Apex Ltd (AL) is attempting to determine whether three production methods (Method 1 to Method 3) differ in production rates. The following table shows the output by each method over a period of time (2 weeks).

Method 1	408	380	425	400	427
Method 2	415	485	450	420	530
Method 3	385	250	295	402	268

State an appropriate hypothesis and using an appropriate non-parametric method, test the same.

Q1 (b) Gross revenue (GR) for a company as a function of investment in TV advertisement (TVA) and newspaper advertisement (NPA) is predicted as:

GR = 83.2+ 2.29 TVA+ 1.30 NPA

The values of the SST and SSR are 25.5 and 23.435 respectively. (Note: number of observations =8)

Find and interpret R^2 and adjusted value R_a^2

When TVA was the only independent variable, R^2 was 0.653 and R_a^2 was 0.595. Do you still prefer multiple regression? Why or why not?

Q1 (c) A manufacturer of garment claims that 91 % of the products are of the "Best" quality, 8 % are "Seconds" and only 1 % are "defective". To test this claim, 500 garments are selected at random. This results in 434 as "Best", 48 as "Seconds", and remaining as defectives.
Can we reject the manufacturer's claim?   *[Marks: 8+6+6]*

Q2 (a) BDD bank is known for giving credit card facility to its valued customers. The management of BDD bank has established a prior probability of 0.05 that any particular cardholder will default. The bank has further found that the probability of missing one or more monthly payments is 0.20 for customers who do not default. Of course, the probability of missing one or more payments for those who default is 1.

- Given that Mr Ashok has missed a monthly payment, compute the posterior probability that Mr Ashok will default.
- The bank would like to recall its card if the probability that a customer will default is more than 0.20. Should the bank recall its card if the customer misses a monthly payment? Why or why not?

Q2 (b) In each lot (of size N=4), a sample of 'n' units (n=1) is taken. If the number of defectives in this sample is zero (i.e. 'c' =0), then the entire lot of N items is accepted. Else the lot is rejected. Draw a flow chart to explain the sampling scheme and also draw the OC curve for this. For what type of products will this type of sampling procedure be required?

Q2 (c) We observe n=85 values of a random variable X that is thought to have a Poisson distribution, obtaining:

X	0	1	2	3	4	5
Frequency	41	29	9	4	1	1

Is there any reason to suspect Poisson distribution for the above data?   *[Marks: 6+8+6]*

Q3 (a) In a simple linear regression equation $Y_i = \beta_0 + \beta_1 X_i$ +Error, with usual assumptions, we compute from n=16 observations
$\beta_1 = 0.35$;   MSE (Mean Sum Squares of Errors) = 2.3 and $\sum (x_i - Xbar)^2 = 100.0$,
Construct a 95 % Confidence Interval for β_1.

Please Turn Overleaf

Q3 (b) A box contains 8 good items and 2 defective items from a factory's production line. A manager selects four items at random from the box, without replacement. Let X = (the number of defective items in the random sample).
Show that the probability distribution of X is not binomial

Q3 (c) A study done by Indian Airlines measuring the fatigue and stress of Air Traffic controllers (ATCs) has resulted in proposals for modifications and redesign of the ATC's workstation. Three such alternative systems (S1, S2 and S3) have been selected to control the stress. The following is a summary of stress data (on some scale) for the ATCs under these systems

Replication	System		
	S1	S2	S3
1	15	15	18
2	14	14	14
3	10	11	15
4	13	12	17
5	16	13	16
6	13	13	13

Use an appropriate ANOVA, state and test the hypothesis and advise Indian Airlines.

[Marks: 6+4+10]

Q 4(a) The electric light bulbs of brand A have a mean lifetime of 1400 hours with a standard deviation of 200 hours , while those of brand B have a mean life of 1200 hours with a standard deviation of 100 hours. If random samples of 125 bulbs of each brand are tested, what is the probability that the brand A bulbs will have a mean lifetime which is at least 160 hours more than the brand B bulbs?

Q4 (b) Explain with an example, the role of statistics in conducting a simulation project.

Q4 (c) In a one-factor experiment with four treatments, the following results were obtained. Construct the ANOVA table and test whether there are significant treatment differences.

Treatment	A	B	C	D
Sample size(n)	20	20	18	18
Sample Mean	40.2	38.6	43.5	50.0
sample variance	900	800	960	720

Assume $F_{table\ at\ 5\%}(3, 72) = 2.74$

[Marks: 8+4+8]

Q5 (a) The knowledge of basic statistical concepts of 10 engineers selected at random was measured on a scale of 0 to 100, before and after a short term course on six-sigma. The results are given below:

Engineer	1	2	3	4	5	6	7	8	9	10
Score before	43	82	77	39	51	66	55	61	79	43
Score after	51	84	74	48	53	61	59	75	82	48

Formulate and test an appropriate hypothesis. (Do not use a *non-parametric* test here!)

Q5 (b) Consider a metal casting process, where the response (Hardness number) is supposedly dependant on Factors such as % of Copper and Magnesium content. Each factor is varied at two levels ("Low" and "High"). Each trial was run twice and the response was measured. The results of this experimentation are given below. Perform ANOVA on the following results and analyze these results in as much detail as possible.

Trial No	Factor A % of Copper content	Factor B % of Magnesium content	Interaction A x B	Response (Hardness Number)	
1	Low	Low	Low	6	8
2	Low	High	High	7	8
3	High	Low	High	3	4
4	High	High	Low	9	10

Explain how the above methodology can be used in a six-sigma project. **[Marks: 8+12]**

PAPER ENDS