(3 Hours) **Total Marks: 80**

N.B. : **(1)** Question No. 1 is compulsory.

(2) Attempt any three questions out of remaining five questions

Q.1. (a) Find the eigen values and eigen vectors of $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ **(5)**

(b) A random variable X has following probability distribution **(5)**

X	1	2	3	4	5	6	7
P(X = x)	K	2K	3K	K^2	K^2 +K	$2K^2$	$4K^2$

Find (i) Value of K and Mean of X

(ii) P (X < 5)

(iii) P (X > 5)

(c) Compare discrete and continuous data. **(5)**

(d) Obtain the Hessian Matrix for the function **(5)**

$$Z = 100 - 10x_2 - 8x_1 - 12x_3 + x_1^2 + x_2^2 + x_3^2$$

Q.2. (a) Find Singular Value of Decomposition of matrix $A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}$ **(10)**

(b) An investigation into the equality of standard deviation of two normal populations gave the following results. **(10)**

Sample	**Size**	**Sample Mean**	**Sum of squares of deviations from mean**
1	13	18	100
2	21	24	150

Test the equality of sample variances at 5% level of significance.
(Given $F_{0.025}$ = 2.68 for degrees of freedom 12 and 20 and $F_{0.025}$ = 3.07 for degrees of freedom 20 and 12.

Q.3. (a) A certain drug is claimed to be effective in curing fever in an experiment on 164 persons with fever. Half of them were given drug and half were given sugar pills. The results obtained are shown in the following table. Test the hypothesis that the drug is effective in curing fever. **(10)**

	Helped	**Harmed**	**No Effect**
Drug	52	10	20
Sugar Pills	44	12	26

(Table value at 5% level of significance for degrees of freedom = 2 is 5.99)

(b) What is a scatter plot and explain types of correlation in scatter plot with example. **(10)**

Q.4. (a) Find 3 yearly moving averages and represent these on a graph paper. Also represent the original time series on the graph. **(10)**

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
Sales (in lakhs)	8	10	12	11	13	15	14	17	17

(b) Find a root of an equation $f(x) = x^3 - 9x - 1$ using False position method in the interval (1, 2) upto four places of decimals. **(10)**

Q.5. (a) Use steepest decent method for $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ starting from the point $X_1 = (0, 0)$. **(10)**

(b) Explain the need for exploratory data analysis. Also list and explain exploratory data analysis techniques. **(10)**

Q.6. (a) Describe with example and action to be taken for the following **(10)**

1. Data Cleaning
2. Irrelevant data
3. Incorrect data
4. Handle Missing Data
5. Outliers

(b) Write a short note on linear discriminant analysis techniques and principal component analysis algorithm **(10)**
