

Statistical Modelling for Machine Learning (313307)

MSBTE Diploma · K-Scheme · Model Question Paper — Set B

Time: 3 Hours

Max. Marks: 70

Instructions:

1. All questions are compulsory.
2. Answer each next main question on a new page.
3. Illustrate your answers with neat sketches wherever necessary.
4. Figures to the right indicate full marks.
5. Assume suitable data, if necessary.
6. Use of non-programmable electronic pocket calculator is permissible.
7. Mobile phone, pager and any other electronic communication devices are not permissible in the examination hall.

Q.1 Attempt any FIVE of the following:

10

- a) The mean marks of 60 students of division A are 55 and the mean marks of 40 students of division B are 65. Find the combined mean marks of all 100 students.
- b) Find the sixth decile (D_6) of the following data: 34, 21, 45, 29, 38, 52, 26, 41, 33, 48, 30, 37.
- c) Find the covariance between x and y:

x	2	4	6	8	10
y	5	9	10	14	17

- d) If the coefficient of correlation between x and y is 0.8 and the regression coefficient of y on x is 1.6, find the regression coefficient of x on y.
- e) Three unbiased coins are tossed together. Find the probability of getting at least two heads.
- f) Construct the backward difference table for the following data and write the values of $\nabla^2 y_5$ and $\nabla^3 y_5$:

x	1	2	3	4	5
y	4	12	32	70	132

- g) Define Type I error and Type II error.

Q.2 Attempt any THREE of the following:

12

- a) Draw a less-than ogive for the following data and locate the median graphically. Verify it by formula.

Marks	0–10	10–20	20–30	30–40	40–50	50–60
Frequency	4	10	18	22	12	4

- b) Draw a histogram for the following distribution and find the mode graphically. Verify it by formula.

Class	20–30	30–40	40–50	50–60	60–70	70–80
Frequency	8	15	23	18	10	6

- c) The runs scored by two batsmen A and B in six innings are given below. Find which batsman is more consistent.

Batsman A	42	38	55	61	35	49
Batsman B	45	50	47	52	43	49

- d) For a distribution, Karl Pearson's coefficient of skewness (based on median) is 0.42, mean = 36 and median = 34.6. Find the standard deviation and the coefficient of variation.

Q.3 Attempt any THREE of the following:

12

- a) Calculate Bowley's coefficient of skewness for the following distribution of weekly wages (₹ hundred):

Wages	10–20	20–30	30–40	40–50	50–60	60–70
Frequency	6	14	25	30	15	10

- b) Calculate Karl Pearson's coefficient of skewness for the following data:

x	5	10	15	20	25	30
f	4	9	16	12	6	3

- c) Calculate Spearman's rank correlation coefficient between the marks given by two examiners to 8 students:

Examiner I	70	65	80	65	90	75	65	85
Examiner II	60	58	72	70	85	66	62	80

- d) Fit a straight-line trend to the following sales data (₹ lakh) by the method of least squares and estimate the sales for 2026:

Year	2018	2019	2020	2021	2022	2023	2024
Sales	12	15	14	18	21	20	25

Q.4 Attempt any THREE of the following:

12

- a) For 8 pairs of observations: $\Sigma x = 40$, $\Sigma y = 56$, $\Sigma x^2 = 240$, $\Sigma y^2 = 436$, $\Sigma xy = 310$. Find the regression line of y on x , estimate y when $x = 7$, and find the coefficient of correlation.
- b) Calculate Karl Pearson's coefficient of correlation between the price (x) and demand (y) of a product and interpret the result:

Price (x)	15	18	21	24	27	30
Demand (y)	52	48	45	40	38	33

- c) If $P(A) = 0.5$, $P(B) = 0.4$ and $P(A \cap B) = 0.2$, find (i) $P(A \cup B)$, (ii) $P(A' \cap B')$, (iii) $P(A | B)$. (iv) Are A and B independent?
- d) A box contains 6 red, 4 green and 5 blue balls. Three balls are drawn at random. Find the probability that (i) all three are of different colours, (ii) all three are red.
- e) Bag I contains 3 white and 5 black balls and Bag II contains 6 white and 2 black balls. A fair coin is tossed: if head, a ball is drawn from Bag I, otherwise from Bag II. The ball drawn is white. Find the probability that it came from Bag II.

Q.5 Attempt any TWO of the following:

12

- a) (i) Using Lagrange's interpolation formula, find $f(4)$ from:

x	0	2	3	6
f(x)	-4	2	14	158

(ii) A curve passes through the points $(0, 1)$, $(1, 4)$, $(2, 11)$, $(3, 22)$ and $(4, 37)$. Using Newton's forward interpolation formula, find the equation $y = f(x)$ of the curve.

- b) (i) The square roots of some numbers are given below. Find $\sqrt{38}$ using Newton's backward interpolation formula.

x	20	25	30	35	40
$\sqrt{x}$	4.4721	5.0000	5.4772	5.9161	6.3246

(ii) Taking $h = 1$, evaluate $\Delta(x^3)$ and $\Delta^2(x^3)$.

- c) The cubes of some integers are given below. Using extrapolation (Newton's forward formula), find $(8)^3$ and $(2)^3$.

x	3	4	5	6
$y = x^3$	27	64	125	216

Q.6 Attempt any TWO of the following:

12

- a) Four coins were tossed 160 times and the number of heads was recorded as below. Test at 5% level of significance whether the coins are unbiased (fit a binomial distribution). (χ^2 for 4 d.f. at 5% = 9.49)

No. of heads	0	1	2	3	4
Frequency	8	34	69	43	6

- b) (i) In city A, 480 out of 800 persons and in city B, 420 out of 600 persons were found to use online payment. Is there a significant difference between the two proportions at 5% level?
(ii) A sample of 64 bulbs has mean life 1180 hours with standard deviation 96 hours. Can the claim that the mean life of the bulbs is 1200 hours be accepted at 5% level of significance?
- c) Weights (kg) of 8 persons before and after a diet programme are given below. Test whether the diet programme is effective in reducing weight. (t for 7 d.f. at 5% = 2.365)

Person	1	2	3	4	5	6	7	8
Before	72	68	80	75	90	66	84	78
After	70	65	78	74	85	66	80	75

Note: This is a **model question paper prepared by Ananda Technologies** only for practice and exam preparation, based on the pattern of previous MSBTE question papers and the K-scheme syllabus. It is **not an official MSBTE paper**, and it does **not** mean that these questions will come in the exam. Answers follow standard textbook points; marks shown are indicative. Free for students — anandatechnologies.in