

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

May-June 2025		
(B. Tech.) Program: Computer/EXTC/IT Engineering Scheme II		
Regular Examination: TY Semester: VI		
Course Code: HDSC601 and Course Name: Statistical Learning for Data Science		
Date of Exam: 30 May 2026	Duration: 02.5 Hours	Max. Marks: 60

Instructions: (1)All questions are compulsory. (2)Draw neat diagrams wherever applicable. (3)Assume suitable data, if necessary.																						
Q. No.	Question						Max. Marks	CO	BT level													
Q 1	Solve any two questions out of three: (05 marks each)						10															
a)	Explain Median and Mode with example?							CO1	U													
b)	Explain Non-parametric Methods with example ?							CO3	U													
c)	Find a least square straight line for the following data? <table><tr><td>X:</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr><tr><td>Y</td><td>6</td><td>4</td><td>3</td><td>5</td><td>4</td><td>2</td></tr></table> And estimate (predict) Y at X=4 and X at Y=4.							X:	1	2	3	4	5	6	Y	6	4	3	5	4	2	CO5
X:	1	2	3	4	5	6																
Y	6	4	3	5	4	2																
Q 2	Solve any two questions out of three: (05 marks each)						10															
a)	Explain about the Errors in Testing of Hypothesis?							CO4	U													
b)	A ball is drawn at random from a box containing 12 red, 18 white, 10 blue and 15 orange balls. Find the probability that (i) It is red or blue (ii) white ,blue or orange (iii) Neither white nor orange							CO2	U													
c)	Write a short note on forecast accuracy.							CO5	U													
Q.3	Solve any two questions out of three. (10 marks each)						20															
a)	Illustrate on cross tabulation and scatter diagram to statistical data science with suitable example?							CO1	U													
b)	A random variable X has the following probability function							CO2	AP													

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

May-June 2026		
(B. Tech.) Program: Computer/EXTC/IT Engineering Scheme II		
Regular Examination: TY Semester: VI		
Course Code: HDSC601 and Course Name: Statistical Learning for Data Science		
Date of Exam: 30 May 2026	Duration: 02.5 Hours	Max. Marks: 60

	<table><tr><td>Values of X:</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr><tr><td>P(x)</td><td>0</td><td>k</td><td>2k</td><td>2k</td><td>3k</td><td>k²</td><td>2 k²</td><td>7 k² + k</td></tr></table> (i)Find k (ii) Evaluate P (X < 6)(iii) P (X ≥ 6) (iv) P (0 < X < 5) (v) If P(X ≤ a) > (1/2) , find the minimum value of a and (vi) Determine the distribution function of X .	Values of X:	0	1	2	3	4	5	6	7	P(x)	0	k	2k	2k	3k	k ²	2 k ²	7 k ² + k								
Values of X:	0	1	2	3	4	5	6	7																			
P(x)	0	k	2k	2k	3k	k ²	2 k ²	7 k ² + k																			
c)	The heights of six randomly chosen sailors are in inches: 63, 65,68,69,71 and 72. The heights of ten randomly chosen soldiers are : 61,62,65,66,69,69,70,71,72 and 73. Discuss in the light that these data throw on the suggestion that soldiers on an average are taller than sailors. (Table value of t for 13 d.f. is 2.16, for 14 d.f. is 2.15 and for 15 d.f. is 2.13)	CO4	AP																								
Q.4	Solve any two questions out of three. (10 marks each)	20																									
a)	A panel of judges A and B graded seven debators and independently awarded the following marks : <table><tr><td>Debator</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr><tr><td>Marks by A</td><td>40</td><td>34</td><td>28</td><td>30</td><td>44</td><td>38</td><td>31</td></tr><tr><td>Marks by B</td><td>32</td><td>19</td><td>26</td><td>30</td><td>38</td><td>34</td><td>28</td></tr></table> An eight debator was awarded 36 marks by Judge A while Judge B was not present. If Judge B was also present, how many marks would you expect him to award to eight debator assuming some degree of relationship exists in judgment?	Debator	1	2	3	4	5	6	7	Marks by A	40	34	28	30	44	38	31	Marks by B	32	19	26	30	38	34	28	CO5	AP
Debator	1	2	3	4	5	6	7																				
Marks by A	40	34	28	30	44	38	31																				
Marks by B	32	19	26	30	38	34	28																				
b)	Explain Systematic sampling, Multistage sampling and Cluster sampling with example?	CO3	U																								
c)	For a certain communication channel the probability that '0'is received as '0' is 0.8. while the probability that '1' is received as '1' is 0.95 .If the probability of transmitting '0' is 0.45 , find (i) a '1' is received. (ii) '1' was transmitted given that '1' was received. (iii) Probability that an error has occurred.	CO2	AP																								
