Name: Enrolment No:	UPES UNIVERSITY WITH A PURPOSE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, December 2021

Course: Probability and Statistics for Engineers — **Semester: III**
Program: B.Tech. (CSE) — **Time : 03 hrs.**
Course Code: CSEG 2036P — **Max. Marks: 100**
No. of printed pages: 3

Instructions: Attempt all the questions. Refer appendix for required distribution tables.

Section A
(Scan and Upload) **(5Qx 4M = 20 Marks)**

S. No.		Marks	CO
Q1	Find the first four moments about the origin for a random variable X having density function $$f(x) = \begin{cases} 4x(9-x^2)/81 & 0 \le x \le 3 \\ 0 & otherwise \end{cases}$$	**[4]**	**CO3**
Q2	A continuous random variable X has probability density given by $$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \le 0 \end{cases}$$ Find (a) $E(X)$ (b) $E(X^2)$	**[4]**	**CO1**
Q3	Discuss Chebyshev's Inequality. A random variable X has mean 3 and variance 2. Use Chebyshev's inequality to obtain an upper bound for $$P(\lvert X-3\rvert \ge 1)$$	**[4]**	**CO2**
Q4	Find the probability that in tossing a fair coin three times, there will appear (a) 3 heads, (b) 2 tails and 1 head	**[4]**	**CO1**
Q5	What is the Decision Tree Algorithm? Discuss the different types of nodes in Decision Trees.	**[4]**	**CO5**

Section B
(Scan and Upload) **(4Qx10M = 40 Marks)**

Q6	Briefly differentiate the following terms: a) Classification and clustering b) Binomial distribution and negative binomial distribution c) Discrete random variable and continuous random variable d) Simple linear regression and multiple regression	**[10]**	**CO2**

<table>
<tr><td>Q7</td><td>For each claim, state H_0 and H_a in words and in symbols. Then determine whether the hypothesis test is a left-tailed test, right-tailed test, or two-tailed test. Sketch a normal sampling distribution and shade the area for the P-value.
a) A school publicizes that the proportion of its students who are involved in at least one extracurricular activity is 61%.
b) A car dealership announces that the mean time for an oil change is less than 15 minutes.
c) A company advertises that the mean life of its furnaces is more than 18 years.</td><td>[10]</td><td>CO3</td></tr>
<tr><td>Q8</td><td>Discuss the characteristics of Bernoulli trial. You are a telemarketer with a 10% chance of persuading a randomly selected person to switch to your long-distance company. You make 8 calls. What is the probability that exactly one is successful?
or
The average playing time of CDs in a large collection is 35 minutes, and the standard deviation is 5 minutes.
a) What value is 1 standard deviation above the mean? 1 standard deviation below the mean? What values are 2 standard deviations from the mean?
b) Assuming the distribution of time is approximately normal, about what percentage of times are between 25 and 45 minutes?</td><td>[10]</td><td>CO3</td></tr>
<tr><td>Q9</td><td>What is meant by Poisson process? Let X equal the number of typos on a printed page with a mean of 3 typos per page.
a) What is the probability that a randomly selected page has at least one typo on it?
b) What is the probability that a randomly selected page has at most one typo on it?</td><td>[10]</td><td>CO2</td></tr>
<tr><td colspan="4">SECTION-C
(Scan and Upload) (2Qx 20M= 40 Marks)</td></tr>
<tr><td>Q10</td><td>A sports statistician claims that the mean winning times for Boston Marathon women’s open division champions is at least 2.68 hours. The mean winning time of a sample of 30 randomly selected Boston Marathon women’s open division champions is 2.60 hours. Assume the population standard deviation is 0.32 hour. (Refer Z table as provided in Appendix-1)
a) Identify the claim and state H_0 and H_a.
b) Find the standardized test statistic z.
c) Find the corresponding P value.
d) At alpha = 0.05, decide whether to reject or fail to reject the null hypothesis.
e) Interpret the decision in the context of the original claim</td><td>[20]</td><td>CO4</td></tr>
<tr><td>Q11</td><td>The number of officers on duty in a Delhi and the number of robberies for that day are:
<table><tr><td>Officers</td><td>10</td><td>15</td><td>16</td><td>1</td><td>4</td><td>6</td><td>18</td><td>12</td><td>14</td><td>7</td></tr><tr><td>Robberies</td><td>5</td><td>2</td><td>1</td><td>9</td><td>7</td><td>8</td><td>1</td><td>5</td><td>3</td><td>6</td></tr></table>Calculate the regression line for this data, and the residual for the first observation, (10; 5). What percentage of variation is explained by the regression line?
OR</td><td>[20]</td><td>CO5</td></tr>
</table>

A study involved comparing the per capita income (in thousands) to the number of medical doctors per 10,000 residents. Six small cities in Uttarakhand had the observations:

Per capita income	8.6	9.3	10.1	8.0	8.3	8.7
Doctors	9.6	18.5	20.9	10.2	11.4	13.1

Calculate the regression line for this data. What percentage of variation is explained by the regression line? Predict the number of doctors per 10,000 residents in a town with a per capita income of 8500.

Appendix 1

Standard Normal Cumulative Probability Table

Cumulative probabilities for NEGATIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379