

Name:

Enrolment No:



**UNIVERSITY OF PETROLEUM AND ENERGY STUDIES**  
**End Semester Examination, April/May 2018**

**Course: Applied Statistical Analysis (CSIB 225)**  
**Program: B.Tech CSE ECRA**  
**Time: 03 hrs.**

**Semester: VI**

**Max. Marks: 100**

**Instructions:**

**SECTION A**

**Note: All 5 questions are compulsory. Each question of Section A carries 4 marks.**

S. No.		Marks	CO
Q 1	Explain the role of statistics in research and development	4	CO1
Q 2	What is ANOVA? State different type of ANOVA with example.	4	CO3
Q 3	What are the four levels of measurements in statistics?	4	CO2
Q 4	Distinguish between Supervised and Unsupervised learning.	4	CO5
Q 5	What is Correlation? What are the different types of correlation?	4	CO4

**SECTION B**

**Note: Answer all the questions. Each question of section B carries 10 marks.**

Q 6	<p>In a population average IQ is 100. A team of scientists wants to test a new medication to see if it has either a positive or negative effect on intelligence or no effect at all. A sample of 30 participants who have taken the medication has a mean of 140 with a standard deviation of 20. Did the medication affect intelligence (Given <math>\alpha=0.05</math>)</p> <p style="text-align: center;">OR</p> <p>Distinguish between</p> <ol style="list-style-type: none"><li>Null hypothesis and alternative hypothesis</li><li>Type I error and Type II error</li></ol>	10	CO3
Q 7	The number of calls coming per minute into a hotels reservation center is Poisson random variable with mean 3.	10	CO1,C O2

	<p>a. Find the probability that no calls come in a given 1 minute period.</p> <p>b. Assume that the number of calls arriving in two different minutes are independent. Find the probability that at least two calls will arrive in a given two minute period.</p>		
Q 8	<p>a. How correlation is different from regression?</p> <p>b. What is the role of data mining technique in statistics?</p>	10	CO4, CO5
Q 9	<p>Write short note on</p> <p>c. Binomial distribution</p> <p>d. Poisson distribution</p>	10	CO2

**SECTION-C**

**Note: Answer the questions. Each question of Section C carries 20 marks.**

Q 10	<p>In a sample survey of public opinion answer to the questions</p> <p>1. Do you drink?</p> <p>2. Are you in favor of local option on sale of liquor tabulated below?</p> <p>Given that the value of chi square for Degree of freedom at 5% level of significance is 3.841</p> <table border="1" data-bbox="162 1045 651 1199"> <thead> <tr> <th></th> <th>YES</th> <th>NO</th> <th>TOTAL</th> </tr> </thead> <tbody> <tr> <th>YES</th> <td>56</td> <td>31</td> <td>87</td> </tr> <tr> <th>NO</th> <td>18</td> <td>6</td> <td>24</td> </tr> <tr> <th>TOTAL</th> <td>74</td> <td>37</td> <td>111</td> </tr> </tbody> </table> <p>Can you infer the local option the local option on the sale of liquor is depend on individual drink?</p> <p align="center">OR</p> <p>Write Short note on</p> <p>a. Decision Tree</p> <p>b. Linear Regression</p> <p>c. Support Vector Machine</p> <p>d. Artificial Neural Network</p>		YES	NO	TOTAL	YES	56	31	87	NO	18	6	24	TOTAL	74	37	111	20	CO1, CO3, CO5		
	YES	NO	TOTAL																		
YES	56	31	87																		
NO	18	6	24																		
TOTAL	74	37	111																		
Q 11	<p>Calculate coefficient between X and Y using Karl Pearson coefficient correlation</p> <table border="1" data-bbox="203 1703 1237 1780"> <tbody> <tr> <td>X</td> <td>23</td> <td>27</td> <td>28</td> <td>28</td> <td>29</td> <td>30</td> <td>31</td> <td>33</td> </tr> <tr> <td>Y</td> <td>18</td> <td>20</td> <td>20</td> <td>27</td> <td>21</td> <td>29</td> <td>27</td> <td>29</td> </tr> </tbody> </table>	X	23	27	28	28	29	30	31	33	Y	18	20	20	27	21	29	27	29	20	CO4
X	23	27	28	28	29	30	31	33													
Y	18	20	20	27	21	29	27	29													