U.S.N. | | | | | | | | | |

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

August 2024 Semester End Main Examinations

Programme: B.E. **Semester: IV**
Branch: Industrial Engineering and Management **Duration: 3 hrs.**
Course Code: 23IM4BSSFE **Max Marks: 100**
Course: Statistics for Engineers

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

UNIT – I

CO *PO* **Marks**

1 a) Explain in brief the steps involved in engineering methods of problem solving with an example. — ***CO** 1* ***PO1*** **06**

b) The ages of a sample of the students attending college this semester are: — ***CO 2*** ***PO1*** **08**

19	17	15	20	23	41	33	21	18	20
18	33	32	29	24	19	18	20	17	22
55	19	22	25	28	30	44	19	20	39

i. Construct a frequency distribution with intervals 15-19, 20-24, 25-29, 30-34, and 35 and older.
ii. Estimate the modal value
iii. Now compute the mean for raw data.

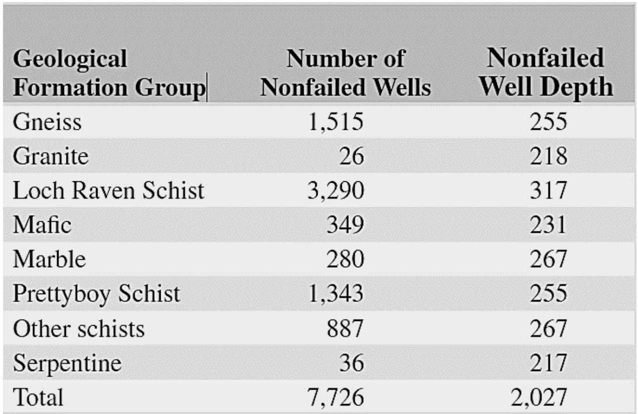
c) Suppose we want to analyse the data of defects on four types of defects complained by the customers, following are the data received information. — ***CO 2*** ***PO1*** **06**

i. Construct the Pareto Chart for the defects complained
ii. Mention the vital few factors and useful many factors affecting the problem

Types of Defects	**Frequency**
Button defect	23
Pocket defect	16
Collar defect	10
Cuff defect	7
Sleeve defect	3
Total	**59**

UNIT – II

2 a) The following table shows the typical depth (rounded to the nearest foot) for nonfailed wells in geological formations in Baltimore County (The Journal of Data Science, 2009, Vol.7, pp. 111–127). — ***CO 2*** ***CO 3*** ***PO1*** ***PO2*** **06**



Geological Formation Group	Number of Nonfailed Wells	Nonfailed Well Depth
Gneiss	1,515	255
Granite	26	218
Loch Raven Schist	3,290	317
Mafic	349	231
Marble	280	267
Prettyboy Schist	1,343	255
Other schists	887	267
Serpentine	36	217
Total	7,726	2,027

(i) Calculate the probability mass function of depth for nonfailed wells from the table.
(ii) Determine the cumulative distribution function for the random variable
(iii) Calculate the mean and variance for the random variable

	b)	The number of flaws in bolts of cloth in textile manufacturing is assumed to be Poisson distributed with a mean of 0.1 flaw per square meter. (i) What is the probability that there are two flaws in one square meter of cloth? (ii) What is the probability that there is one flaw in 10 square meters of cloth? (iii) What is the probability that there are no flaws in 20 square meters of cloth? (iv) What is the probability that there are at least two flaws in 10 square meters of cloth?	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**08**
	c)	Problem illustrates that poor quality can affect schedules and costs. A manufacturing process has 100 customer orders to fill. Each order requires one component part that is purchased from a supplier. However, typically, 2% of the components are identified as defective, and the components can be assumed to be independent. (i) If the manufacturer stocks 100 components, what is the probability that the 100 orders can be filled without reordering components? (ii) If the manufacturer stocks 102 components, what is the probability that the 100 orders can be filled without reordering components? (iii) If the manufacturer stocks 105 components, what is the probability that the 100 orders can be filled without reordering components?	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**06**
		OR			
3	a)	The probability density function of the length of a metal rod is f (x) = 2 for 2.3 < x < 2.8 meters. (i) If the specifications for this process are from 2.25 to 2.75 meters, what proportion of rods fail to meet the specifications?	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**06**

			(ii) Assume that the probability density function is f (x) = 2 for an interval of length 0.5 meters. Over what value should the density be centered to achieve the greatest proportion of rods within specifications? (iii) Determine the cumulative distribution function for the distribution. Use the cumulative distribution function to determine the probability that a length exceeds 2.7 meters.			
	b)	The life of a semiconductor laser at a constant power is normally distributed with a mean of 7000 hours and a standard deviation of 600 hours. (i) What is the probability that a laser fails before 5000 hours? (ii) What is the life in hours that 95% of the lasers exceed? (iii) If three lasers are used in a product and they are assumed to fail independently, what is the probability that all three are still operating after 7000 hours?	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**06**
	c)	The number of (large) inclusions in cast iron follows a Poisson distribution with a mean of 2.5 per cubic millimeter. Approximate the following probabilities: (i) Determine the mean and standard deviation of the number of inclusions in a cubic centimeter (cc). (ii) Approximate the probability that fewer than 2600 inclusions occur in a cc. (iii) Approximate the probability that more than 2400 inclusions occur in a cc. (iv) Determine the mean number of inclusions per cubic millimeter such that the probability is approximately 0.9 that 500 or fewer inclusions occur in a cc.	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**08**
		UNIT - III			
4	a)	Distinguish between point estimate and interval estimate with an example.	***CO*** 1	***PO1***	**06**
	b)	Of 1000 randomly selected cases of lung cancer, 823 resulted in death within 10 years. (i) Calculate a 95% two-sided confidence interval on the death rate from lung cancer. (ii) Using the point estimate of p obtained from the preliminary sample, what sample size is needed to be 95% confident that the error in estimating the true value of p is less than 0.03? (iii) How large must the sample be if you wish to be at least 95% confident that the error in estimating p is less than 0.03, regardless of the true value of p?	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**06**
	c)	An article in the Australian Journal of Agricultural Research [“Non-Starch Polysaccharides and Broiler Performance on Diets Containing Soyabean Meal as the Sole Protein Concentrate” (1993, Vol. 44(8), pp. 1483–1499)] determined that the essential amino acid (Lysine) composition level of soybean meals is as shown here (g/kg):	***CO 2*** ***CO 3***	***PO1*** ***PO2***	**08**

<table>
<tr><td></td><td></td><td></td><td>
<table>
<tr><td>22.2</td><td>24.7</td><td>20.9</td><td>26.0</td><td>27.0</td></tr>
<tr><td>24.8</td><td>26.5</td><td>23.8</td><td>25.6</td><td>23.9</td></tr>
</table>
(i) Construct a 99% two-sided confidence interval for σ^2.
(ii) Calculate a 99% lower confidence bound for σ^2.
(iii) Calculate a 90% lower confidence bound for σ.
(iv) Compare the intervals that you have computed.
</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td>UNIT – IV</td><td></td><td></td><td></td></tr>
<tr><td></td><td>5</td><td>a)</td><td>Cloud seeding has been studied for many decades as a weather modification procedure (for an interesting study of this subject, see the article in Technometrics, “A Bayesian Analysis of a Multiplicative Treatment Effect in Weather Modification,” Vol. 17,
pp. 161–166). The rainfall in acre-feet from 20 clouds that were selected at random and seeded with silver nitrate follows: 18.0, 30.7, 19.8, 27.1, 22.3, 18.8, 31.8, 23.4, 21.2, 27.9, 31.9, 27.1, 25.0, 24.7, 26.9, 21.8, 29.2, 34.8, 26.7, and 31.6.
(i) Can you support a claim that mean rainfall from seeded clouds exceeds 25 acre-feet? Use α = 0.01. Find the P-value</td><td><i>CO 2
CO 3</i></td><td><i>PO1
PO2</i></td><td>08</td></tr>
<tr><td></td><td></td><td>b)</td><td>Let X denote the number of flaws observed on a large coil of galvanized steel. Of 75 coils inspected, the following data were observed for the values of X:
<table>
<tr><td>Values Observed</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td></tr>
<tr><td>Frequency</td><td>1</td><td>11</td><td>8</td><td>13</td><td>11</td><td>12</td><td>10</td><td>9</td></tr>
</table>
(i) Does the assumption of the Poisson distribution seem appropriate as a probability model for these data? Use α = 0.01.
(ii) Calculate the P-value for this test.</td><td><i>CO 2
CO 3</i></td><td><i>PO1
PO2</i></td><td>12</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>6</td><td>a)</td><td>Consider the hypothesis test $H_0 : \mu 1 = \mu 2$ against $H_1: \mu 1 \neq \mu 2$ with known variances $\sigma_1 = 10$ and $\sigma_2 = 5$. Suppose that sample sizes $n_1 = 10$ and $n_2 = 15$ and that $\bar{x}_1 = 4.7$ and $\bar{x}_2 = 7.8$. Use α = 0.05.
(i) Test the hypothesis and find the P-value.
(ii) Explain how the test could be conducted with a confidence interval.</td><td><i>CO 2
CO 3</i></td><td><i>PO1
PO2</i></td><td>10</td></tr>
<tr><td></td><td></td><td>b)</td><td>In the 2004 presidential election, exit polls from the critical state of Ohio provided the following results: For respondents with college degrees, 53% voted for Bush and 46% voted for Kerry. There were 2020 respondents.
(i) Is there a significant difference in these proportions? Use α = 0.05. What is the P-value?
(ii) Calculate a 95% confidence interval for the difference in the two proportions and comment on the use of this interval to answer the question in part (i).</td><td><i>CO 2
CO 3</i></td><td><i>PO1
PO2</i></td><td>10</td></tr>
</table>

<table>
<tr><td></td><td></td><td></td><td>UNIT – V</td><td></td><td></td><td></td></tr>
<tr><td></td><td>7</td><td>a)</td><td>A study was made to model the relation between weekly advertising expenditures and sales. During the study following data were recorded:
<table>
<tr><td>Advertising cost (Rs)</td><td>20</td><td>25</td><td>30</td><td>35</td><td>40</td><td>45</td><td>50</td><td>55</td><td>60</td></tr>
<tr><td>Weekly sales</td><td>400</td><td>420</td><td>405</td><td>480</td><td>475</td><td>490</td><td>525</td><td>560</td><td>515</td></tr>
</table>
i. Plot a scatter diagram
ii. Find the equation of the regression line to predict weekly sales from advertising expenditures.
iii. Compute Coefficient of determination R^2 and interpret about model.
iv. Test the hypothesis for Slope using $\alpha = 0.05$</td><td><i>CO 2</i>
<i>CO 3</i></td><td><i>PO1</i>
<i>PO2</i></td><td>16</td></tr>
<tr><td></td><td></td><td>b)</td><td>Discuss briefly the multi linear regression. With its applications.</td><td><i>CO</i> 1</td><td><i>PO1</i></td><td>04</td></tr>
</table>

****