U.S.N.										

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

Autonomous Institute Affiliated to VTU

February 2025 Semester End Main Examinations

Programme: B.E. **Semester: IV**
Branch: Industrial Engineering and Management **Duration: 3 hrs.**
Course Code: 23IM4BSSFE / 22IM4BSSFE **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 the phases of statistics with an example.	CO1	PO1	**04**
	b)	A highly classified study was conducted and the following data was recorded for a sample of 30 observations as shown in the table below 5.02 4.50 4.61 5.55 2.05 2.11 5.51 2.08 2.29 5.46 3.75 5.05 1.50 4.49 3.04 3.35 5.57 3.89 1.67 1.68 3.74 1.67 4.41 2.92 1.76 3.21 5.64 4.03 2.06 6.00 i. Construct box plot for frequencies and comment on outliers. ii. Construct the frequency polygon curve and Ogive curve for the following data iii. Estimate the mean, median, mode and variance.	CO1 CO2	PO1 PO2	**16**
		OR			
2	a)	Explain how mechanistic and Empirical models play an important role in Statistical Modeling Frame work	CO1	PO1	**08**
	b)	The following are the measurements of the product chart in mm 146 148 146 140 150 146 142 150 150 152 146 150 158 148 140 152 160 162 146 156 150 150 146 152 150 148 148 150 150 148 152 146 156 156 142 146 146 154 152 150 I. Estimate the mean, mode, median, standard deviation and variance for the same. II. Construct a frequency polygon and histogram considering all classes. Comment about the population.	CO1 CO2	PO1 PO2	**12**

			UNIT – II			
	3	a)	The number of telephone calls that arrive at a phone exchange is often modeled as a Poisson random variable. Assume that on the average there are 10 calls per hour. i) What is the probability that there are exactly 5 calls in one hour? ii) What is the probability that there are 3 or fewer calls in one hour? iii) What is the probability that there are exactly 15 calls in two hours? iv) What is the probability that there are exactly 5 calls in 30 minutes?	*CO1*	*PO1*	**08**
		b)	Heart failure is due to either natural occurrences (87%) or outside factors (13%). Outside factors are related to induced substances or foreign objects. Natural occurrences are caused by arterial blockage, disease, and infection. Suppose that 20 patients will visit an emergency room with heart failure. Assume that causes of heart failure for the individuals are independent. i. What is the probability that three individuals have conditions caused by outside factors? ii. What is the probability that three or more individuals have conditions caused by outside factors? iii. What are the mean and standard deviation of the number of individuals with conditions caused by outside factors?	*CO1*	*PO1*	**06**
		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. (a) If the manufacturer stocks 100 components, what is the probability that the 100 orders can be filled without reordering components? (b) If the manufacturer stocks 102 components, what is the probability that the 100 orders can be filled without reordering components? (c) If the manufacturer stocks 105 components, what is the probability that the 100 orders can be filled without reordering components?	*CO1* *CO2*	*PO1* *PO2*	**06**
			OR			
	4	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? (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.	*CO1*	*PO1*	**06**

		b)	A study by Bechtel et al., 2009, described in the Archives of Environmental & Occupational Health considered polycyclic aromatic hydrocarbons and immune system function in beef cattle. Some cattle were near major oil- and gas-producing areas of western Canada. The mean monthly exposure to PM1.0 (particulate matter that is < 1μm in diameter) was approximately 7.1 μg/m3with standard deviation 1.5. Assume that the monthly exposure is normally distributed. (i) What is the probability of a monthly exposure greater than 9 μg/m3? (ii) What is the probability of a monthly exposure between 3 and 8 μg/m3? (iii) What is the monthly exposure level that is exceeded with probability 0.05? (iv) What value of mean monthly exposure is needed so that the probability of a monthly exposure more than 9 μg/m3 is 0.01?	*CO1*	*PO1*	**08**
		c)	Suppose that the time it takes a data collection operator to fill out an electronic form for a database is uniformly between 1.5 and 2.2 minutes. (i) What are the mean and variance of the time it takes an operator to fill out the form? (ii) What is the probability that it will take less than two minutes to fill out the form?	*CO1*	*PO1*	**06**
			UNIT - III			
	5	a)	Differentiate between point estimate and interval estimate with an example	*CO1* *CO2*	*PO1* *PO2*	**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?	*CO1*	*PO1*	**06**
		c)	A machine produces metal rods used in an automobile suspension system. A random sample of 15 rods is selected, and the diameter is measured. The resulting data (in millimeters) are as follows: 8.24 8.25 8.20 8.23 8.24 8.21 8.26 8.26 8.20 8.25 8.23 8.23 8.19 8.28 8.24 i) Calculate a 95% two-sided confidence interval on mean rod diameter. ii) Calculate a 95% upper confidence bound on the mean. Compare this bound with the upper bound of the two-sided confidence interval and discuss why they are different.	*CO1* *CO2*	*PO1* *PO2*	**08**
			OR			

	6	a)	A civil engineer is analyzing the compressive strength of concrete. Compressive strength is normally distributed with $\sigma^2 = 1000(psi)^2$. A random sample of 12 specimens has a mean compressive strength of $\overline{x} = 3250$ psi. i. Construct a 95% two-sided confidence interval on mean compressive strength. ii. Construct a 99% two-sided confidence interval on mean compressive strength. Compare the width of this confidence interval with the width of the one found in part (I).	*CO1* *CO2*	*PO1* *PO2*	**06**
		b)	The concentration of active ingredient in a liquid laundry detergent is thought to be affected by the type of catalyst used in the process. The standard deviation of active concentration is known to be 3 grams per liter regardless of the catalyst type. Ten observations on concentration are taken with each catalyst, and the data follow: Catalyst 1: 57.9, 66.2, 65.4, 65.4, 65.2, 62.6, 67.6, 63.7, 67.2, 71.0 Catalyst 2: 66.4, 71.7, 70.3, 69.3, 64.8, 69.6, 68.6, 69.4, 65.3, 68.8 (i) Find a 95% confidence interval on the difference in mean active concentrations for the two catalysts. Find the P-value. (ii) If this difference of 5 grams per liter is really important, do you consider the sample sizes used by the experimenter to be adequate? Does the assumption of normality seem reasonable for both samples?	*CO1* *CO2*	*PO1* *PO2*	**08**
		c)	What are the properties of the good estimator	*CO1*	*PO1*	**06**
			UNIT – IV			
	7	a)	The number of calls arriving at a switchboard from noon to 1:00 p.m. during the business days Monday through Friday is monitored for six weeks (i.e., 30 days). Let X be defined as the number of calls during that one-hour period. The relative frequency of calls was recorded and reported as	*CO1* *CO2*	*PO1* *PO2*	**10**

Value relative	5	6	8	9	10
Frequency	0.067	0.067	0.100	0.133	0.200
Value relative	11	12	13	14	15
Frequency	0.133	0.133	0.067	0.033	0.067

i. Does the assumption of a Poisson distribution seem appropriate as a probability model for this data? Use $\alpha = 0.05$.
ii. Calculate the P-value for this test.

		b)	A 1992 article in the Journal of the American Medical Association ("A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich") reported body temperature, gender, and heart rate for a number of subjects. The body temperatures for 25 female subjects follow: 97.8, 97.2, 97.4, 97.6, 97.8, 97.9, 98.0, 98.0, 98.0, 98.1, 98.2, 98.3, 98.3, 98.4, 98.4, 98.4, 98.5, 98.6, 98.6, 98.7, 98.8, 98.8, 98.9, 98.9, and 99.0. i. Test the hypothesis H_0: $\mu = 98.6$ versus H_1: $\mu \neq 98.6$, using $\alpha = 0.05$. Find the P-value. ii. Compute the power of the test if the true mean female body temperature is as low as 98.0. iii. What sample size would be required to detect a true mean female body temperature as low as 98.2 if you wanted the power of the test to be at least 0.9?	*CO1* *CO2*	*PO1* *PO2*	**10**
			OR			
	8	a)	Two different types of injection-molding machines are used to form plastic parts. A part is considered defective if it has excessive shrinkage or is discolored. Two random samples, each of size 300, are selected, and 15 defective parts are found in the sample from machine 1, and 8 defective parts are found in the sample from machine 2. i) Is it reasonable to conclude that both machines produce the same fraction of defective parts, using $\alpha = 0.05$? Find the P-value for this test. ii) Suppose that $p_1 = 0.05$ and $p_2 = 0.01$. With the sample sizes given here, what is the power of the test for this two sided alternate?	*CO1*	*PO1*	**12**
		b)	In a random sample of 85 automobile engine crankshaft bearings, 10 have a surface finish roughness that exceeds the specifications. Do these data present strong evidence that the proportion of crankshaft bearings exhibiting excess surface roughness exceeds 0.10? i) State and test the appropriate hypotheses using $\alpha = 0.05$.	*CO1*	*PO1*	**08**

<table>
<tr><td></td><td></td><td></td><td>UNIT – V</td><td></td><td></td><td></td></tr>
<tr><td></td><td>9</td><td>a)</td><td>An article in Concrete Research [“Near Surface Characteristics of Concrete: Intrinsic Permeability” (1989, Vol. 41)] presented data on compressive strength x and intrinsic permeability y of various concrete mixes and cures. Summary quantities are n = 14, Σy_i = 572, Σy_i^2 = 23,530, Σx_i = 43, Σx_i^2 = 157.42, and $\Sigma x_i y_i$ = 1697.80. Assume that the two variables are related according to the simple linear regression model.
i) Calculate the least squares estimates of the slope and intercept. Estimate σ^2. Graph the regression line.
ii) Use the equation of the fitted line to predict what permeability would be observed when the compressive strength is x = 4.3
iii) Give a point estimate of the mean permeability when compressive strength is x = 3.7
iv) Suppose that the observed value of permeability at x = 3.7 is y = 46.1. Calculate the value of the corresponding residual</td><td>CO2
CO3</td><td>PO2
PO12</td><td>12</td></tr>
<tr><td></td><td></td><td>b)</td><td>Discuss briefly the Multi linear Regression. With its application.</td><td>CO1</td><td>PO1</td><td>08</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>10</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 α = 0.05
v.</td><td>CO2
CO3</td><td>PO2
PO12</td><td>16</td></tr>
<tr><td></td><td></td><td>b)</td><td>Distinguish between correlation and regression.</td><td>CO1</td><td>PO1</td><td>04</td></tr>
</table>

****