U.S.N.										

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

Autonomous Institute Affiliated to VTU

June 2025 Semester End Main Examinations

Programme: B.E. **Semester: IV**
Branch: Biotechonology **Duration: 3 hrs.**
Course Code: 23MA4BSBDE / 22MA4BSBDE **Max Marks: 100**
Course: Biostatistics and Design of Experiments

Instructions:
1. Each unit has an internal choice, answer one complete question from each unit.
2. Missing data, if any, may be suitably assumed.
3. Use of statistical tables are allowed.

		UNIT - 1	CO	PO	Marks
1	a)	In a chemical engineering experiment, researchers are studying the relationship between **temperature** and the **reaction rate** of a certain process. The following data represent measurements of **temperature (in °C)** and the corresponding **reaction rate (in mol/min)** for 10 trials: Temperature(x): 22.1, 25.3, 28.7, 30.0, 32.5, 35.1, 37.4, 39.6, 42, 45.2 Reaction rate(y): 1.2, 1.5, 2.0, 2.4, 2.8, 3.1, 3.5, 3.9, 4.2, 4.6 Find the correlation between temperature and reaction rate. Hence interpret the result.	1	2	**6**
	b)	An investigator has reported the data tabulated for an experiment to determine the growth rate of bacteria y as a function of oxygen concentration $x(mg/L)$. It is known that such data can be modeled by the equation $y = ab^x$. Fit the equation and predict the growth rate at $x = 5mg/L$. x: 2, 4, 6, 8, 10, 12 y: 1.8, 1.5, 1.4, 1.1, 1.1, 0.9	1	2	**7**
	c)	The lifetime (X) (in years) of a certain type of electronic component is modeled by the probability density function: $f(x) = kx^3,\ 0 < x < 2$. (i) Find the value of the constant k so that $f(x)$ is a valid probability density function. (ii) Compute the mean and variance of X. (iii) What is the probability that a component lasts more than 1 year?	1	2	**7**
		OR			
2	a)	If X is Poisson variate with probability mass function $p(x)$ then prove that the mean and variance of a Poisson distribution are same.	1	1	**6**

		b)	Given $8x-10y+66=0$ and $40x-18y=214$ are the two regression lines. Find (i) mean of *x*-series, *y*-series. (ii) correlation coefficient between *x* and *y*. (iii) σ_y if $\sigma_x=3$.	*1*	*1*	**7**
		c)	The following table gives the measurements of air resistance R (in Newtons) on a vehicle traveling at different speeds V (in km/h): V: 10, 30, 50, 70, 90, 110 R: 2.0, 5.2, 10.1, 17.5, 27.3, 39.8 Assuming the relationship between resistance R and velocity V is quadratic: $R=a+bV+cV^2$. Apply method of least square to find the constants a, b & c.	*1*	*1*	**7**
			UNIT – 2			
3	a)		The manufacturer of a certain medicine claimed that it was 90% effective in relieving an allergy for a period of 8 hours. In a sample of 200 persons who had the allergy, the medicine provided relief for 160 persons. Determine whether the manufacturer's claim is legitimate at 0.01 level of significance.	*1*	*2*	**6**
	b)		Two machines are used to fill plastic bottles with dishwashing detergent. The standard deviations of fill volume are known to be $\sigma_1=0.1$ fluid ounces and $\sigma_2=0.15$ fluid ounces for the two machines, respectively. Two random samples of $n_1=12$ bottles from machine 1 and $n_2=10$ bottles from machine 2 are selected, and the sample mean fill volumes are $\bar{x}_1=30.87$ fluid ounces and $\bar{x}_2=30.68$ fluid ounces. Assuming normality construct a 90% confidence interval on the mean difference in fill volume.	*1*	*2*	**7**
	c)		A test is designed to measure level of anxiety was administered to a sample of male and female patients just prior to undergoing the same surgical procedure. The sample sizes and the variances computed from the scores were as follows: *Male*: $n_1=16$, $S_1^2=150$; *Female*: $n_2=12$, $S_2^2=275$ Apply an F-test at the 5% significance level to determine whether the data provide sufficient evidence that female scores are more variable than male scores in the represented populations.	*1*	*2*	**7**
			OR			
4	a)		The mean life of 100 electric bulbs produced by a company is 2550 hours with a standard deviation of 54 hours. Find 95% confidence limits for population mean life of electric bulbs produced by the company.	*1*	*2*	**6**
	b)		In a winter of an epidemic flu, 2000 babies were surveyed by a well-known pharmaceutical company to determine if the company's new medicine was effective after 2 days. Among 120 babies who had the flu & were given the medicine, 29 were cured within two days. Among 280 babies who had the flu but were not given the medicine, 56 were cured within 2 days. Is there any significant indication that supports the company's claim of the effectiveness of the medicine? Use 5% level of significance.	*1*	*2*	**7**

	c)	In a Mendelian experiment on breeding, four types of plants are expected to occur in the proportion 9:3:3:1. The observed frequencies are 891 round and yellow, 316 wrinkled and yellow, 290 round and green, and 119 wrinkled and green. Apply Chi-square test to test the hypothesis that these data bear out the stated belief. Use 5% level of significance.	*1*	*2*	**7**
		UNIT - 3			
5	a)	A sociologist is interested in comparing the ages of husbands and wives. He collected the data below, which shows the ages for the husband and wife in a random chosen sample of nine couples. Couple: A B C D E F G H I Husband's Age: 79 39 55 71 37 39 48 63 54 Wife's Age: 70 36 49 54 38 32 49 52 56 Apply Wilcoxon signed rank test to test the hypothesis that most men are older than their wives at 5% level of significance.	*1*	*2*	**6**
	b)	The driver of a diesel-powered automobile decided to test the quality of three types of diesel fuel sold in the area based on mileage per litre. Carry out one way ANOVA at 1% level of significance. Brand A: 38.7 39.2 40.1 38.9 -- Brand B: 41.9 42.3 41.3 -- -- Brand C: 40.8 41.2 39.5 38.9 40.3	*1*	*2*	**7**
	c)	Of 18 selected patients who had advance stage of leukemia, 10 were treated with new serum and 8 were not. The survival time was observed from the time of experiment are as follows: Treatment: 2.9 3.1 5.3 4.2 4.5 3.9 2.0 3.7 4.1 4.0 No treatment: 1.9 0.5 0.9 2.2 3.1 2.0 1.7 2.5 - - Apply Mann-Whitney test to test whether the serum is effective at 5% level of significance.	*1*	*2*	**7**
		OR			
6	a)	The following data gives the amount of androgen present in blood of 10 deer's before and 30 minutes after a certain drug is injected to them. *Before*: 2.76 5.18 2.68 3.05 4.10 7.05 6.6 4.79 7.39 7.3 *After*: 7.02 3.1 5.44 3.99 5.21 10.26 13.91 18.53 7.91 4.85 Test at 0.01 level of significance, whether there is significant change in the concentration levels of androgen in blood.	*1*	*1*	**6**

	b)	The effects of two drugs on reaction time to a certain stimulus were studied in three samples of experimental animals. Sample III served as a control while the animals in sample I were treated with drug A and those in sample II were treated with drug B prior to the application of the stimulus. The following table shows the reaction time in seconds of 13 animals. Sample I: 17, 20, 40, 31, 35 Sample II: 8, 7, 9, 8, - Sample III: 2, 5, 4, 3, - Apply Kruskal-Walli's test to conclude whether the three populations represented by the three samples differ with respect to reaction time? Use 1% level of significance.	*1*	*1*	**7**
	c)	A group of 10 rats fed on a diet A and another group of 8 rats fed on a different diet B, recorded the following increase in weights. Diet A: 5, 6, 8, 1, 12, 4, 3, 9, 6 Diet B: 2, 3, 6, 8, 10, 1, 2, 8, -- At the 5% significance level, test whether the mean of diet A is greater than that of diet B.	*1*	*1*	**7**
		UNIT - 4			
7	a)	Explain the different types of scales of measurement used in the descriptive statistics.	*1*	*1*	**6**
	b)	For the data given below, calculate the coefficient of quartile deviation. Marks: 10, 20, 30, 40, 50, 60 No. of Students: 4, 7, 15, 8, 7, 2	*1*	*1*	**7**

	c)	Researchers conduct a 15-year prospective cohort study to investigate the association between low physical activity and the risk of developing heart disease. A total of 500 exposed and 500 unexposed participants are enrolled.
		• Exposed group (low physical activity):
		  ○ Total: 500 participants
		  ○ Developed heart disease: 80
		• Unexposed group (physically active):
		  ○ Total: 500 participants
		  ○ Developed heart disease: 50
		(i) What is the relative risk (RR) of developing heart disease for those with low physical activity?
		(ii) What is the attributable risk?
		(iii) Assume that 40% of the general population is physically inactive (i.e., exposed group). If we could promote physical activity across the population and eliminate low physical activity, what percentage of heart disease cases in the population could be prevented?
			1	*1*	**7**

		OR			
8	a)	Briefly discuss the methodological advantages and limitations of employing a cohort study design.	*1*	*1*	**6**

	b)	Wages of workers of two shifts in a factory is given below. (i) Which shift worker earns better? (ii) Which shift workers are consistent in their earnings.	*1*	*1*	**7**

Wages	50-100	100-150	150-200	200-250	250-300	300-350
Shift A	3	8	24	63	102	50
Shift B	32	44	62	41	37	34

	c)	Explain the key differences between case-control studies and cross-sectional studies in terms of design, purpose and data collection.	*1*	*1*	**7**
		UNIT - 5			
9	a)	An agricultural scientist is interested in an experiment to ascertain the effect of applying chemical fertilizers on the land production in different acres of land. It has divided into four fertilizers treatments farm: A, B, C and D. There are three blocks each containing four acres of land of comparable sizes. As a result of the experiment, the following is the production data from the lands. Apply Randomized Block Design to analyze the data at 1% level of significance.	*1*	*2*	**10**

Blocks (Lands)	Fertilizers			
	A	B	C	D
I	10	12	14	13
II	16	20	24	25
III	18	21	26	28

	b)	An oil company tested 4 different blends - A, B, C and D of gasoline for fuel efficiency according to a Latin Square Design in order to control for the variability of 4 different drivers and 4 different models of cars. Fuel efficiency was measured in miles per gallon after driving the cars over a standard course. The data are presented below. Analyze the data at 5% level of significance.	*1*	*2*	**10**

	CAR MODEL			
Driver	I	II	III	IV
1	D-15.5	B-33.9	C-13.2	A-29.1
2	B-16.3	C-26.6	A-19.4	D-22.8
3	C-10.8	A-31.8	D-17.1	B-30.3
4	A-14.7	D-34	B-19.7	C-21.6

		OR			

	10	a)	Data recorded on yield of four varieties in an experiment with four replications for which one value is missing. Estimate the missing value and analyze the data at 5% level of significance. *P* 5.52 | *R* 5.57 | *Q* 5.071 | *S* 9.16 *S* 6.69 | *R* 5.14 | *Q* – | *P* 6.09 *S* 2.89 | *P* 6.02 | *Q* 6.53 | *R* 2.83 *R* 9.76 | *Q* 6.25 | *S* 8.9 | *P* 9.77	*1*	*2*	**10**
		b)	An experiment was conducted on the yield of potatoes in a randomized block design with four replications. Analyze the following 2^2 factorial design. Use 1% level of significance. Block | Treatment Combinations (1) | (1) 23 | K 25 | P 22 | KP 38 (2) | P 40 | (1) 26 | K 36 | KP 38 (3) | (1) 29 | K 20 | KP 30 | P 20 (4) | KP 34 | K 31 | P 24 | (1) 28	*1*	*2*	**10**

****