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: Biotechnology **Duration: 3 hrs.**
Course Code: 23MA4BSBDE / 22MA4BSBDE **Max Marks: 100**
Course: Biostatistics and Design of Experiments

Instructions:
1. All units have internal choices, answer one complete question from each unit.
2. Missing data, if any, may be suitably assumed.
3. Use of Statistical tables are permitted.

			UNIT - 1	CO	PO	Marks
1	a)		A distributor of bean seeds determines from extensive tests that 5% of large batch of seeds will not germinate. He sells the seeds in packets of 200 and guarantees 98% germination. Determine the probability that a particular packet will violate the guarantee.	1	2	**6**
	b)		A study is conducted involving 7 infants to investigate the association between the gestational age at birth, measured in weeks, and birth weight, measured in grams. Estimate the association between the gestational age and infant birth weight by finding Karl Pearson's coefficient of correlation for the data. gestational age(weeks): 34.7, 36.0, 29.3, 40.1, 35.7, 42.4, 40.3 birth weight(grams): 1895, 2030, 1440, 2835, 3090, 3827, 3260 Hence, predict the birth weight by fitting the appropriate regression line if the gestational age is 38.5.	1	2	**7**
	c)		The following table gives the results of the measurements of train resistances R and the velocity V in miles per hour. V: 20, 40, 60, 80, 100, 120 R: 5.5, 9.1, 14.9, 22.8, 33.3, 46.0 If R is related to V by the relation $R = a + bV + cV^2$, then find a, b and c.	1	2	**7**
			OR			
2	a)		If X is a 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)	The content, in milligrams, of vitamin C in a litre carton of cranberry juice can be modelled by a normal distribution with mean of 32 and standard deviation of 2. i) Determine the probability that, for a carton chosen at random, the vitamin C content is less than 30mg. ii) Determine the vitamin C content to the closest milligram in order to confirm that the percentage of cartons is 2.5.	*1*	*2*	**7**
		c)	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 $x = 5$ mg/L. x: 2, 4, 6, 8, 10, 12 y: 1.8, 1.5, 1.4, 1.1, 1.1, 0.9	*1*	*2*	**7**
			UNIT - 2			
	3	a)	A wholesaler in apples claims that only 4% of the apples supplied by him are rotten. A random sample of 600 apples contained 36 rotten apples. At 5% level of significance, test whether the claim of the wholesaler is violated.	*1*	*2*	**6**
		b)	In a study of nutrition cares in nursing homes, it is found that among 55 patients with hypertension, 24 were on sodium restricted diets. Of 149 patients without hypertension, 36 were on sodium restricted diets. Construct a 99% confidence interval for the difference in the proportion of patients on sodium restricted diets with hypertension and without hypertension?	*1*	*2*	**7**
		c)	Genetic theory states that children having one parent of blood type *M* and other blood type *N* will always one of three types *M*, *MN*, *N* and that the proportions of these types will on average be 1: 2: 1. A report states that out of 300 children having one parent *M* and one *N* parent 30% were found to be of type *M*, 45% of type *MN* and remainder of type *N*. Find the Chi-square value and examine the correspondence between the theory and the experiment. Use 5% level of significance.	*1*	*2*	**7**
			OR			
	4	a)	The mean life of 100 electric blubs produced by a company is 2550 hours with a standard deviation 54 hours. Find 95% confidence limits for population mean life of electric blubs produced by the company.	*1*	*2*	**6**
		b)	In 210 families of females with primary unipolar major depression, they found that alcoholism was present in 89. Of 299 control families, alcoholism was present in 94. At 1% level of significance, do these data provide sufficient evidence for us to conclude that alcoholism is more likely to be present in females of subjects with unipolar depression?	*1*	*2*	**7**

	c)	Two types of instruments for measuring the amount of Sulphur monoxide in the atmosphere are being compared in an air pollution experiment. The following readings were recorded for the two instruments. Instrument A: 0.86, 0.82, 0.75, 0.61, 0.89, 0.64, 0.81, 0.68, 0.65 Instrument B: 0.87, 0.74, 0.63, 0.55, 0.76, 0.7, 0.69, 0.57, 0.53 Apply F-test at 1% level of significance to check whether the two types of instruments yield measurements having the same variability.	*1*	*2*	**7**
		UNIT - 3			
5	a)	Five pre-school children were given a supplement of multipurpose food for a period of four months. Their skin-fold thickness (in mm) was measured before the commencement of the programme and also at the end. Apply the parametric test to check if there is any change in their skin-fold thickness for the following data. Use 5% level of significance. Child: 1, 2, 3, 4, 5 At the beginning of the program: 6, 8, 8, 6, 9 At the end of the training program: 8, 8, 10, 7, 10	*1*	*2*	**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 times 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-Wallis One Way Analysis of Variance to conclude whether the three populations represented by the three samples differ with respect to reaction time? Use 1% level of significance.	*1*	*2*	**7**
	c)	Consider a Phase II clinical trial designed to investigate the effectiveness of a new drug to reduce symptoms of asthma in children. A total of $n = 10$ participants are randomized to receive either the new drug or a placebo. Participants are asked to record the number of episodes of shortness of breath over a 1-week period following receipt of the assigned treatment. The data are shown below. Placebo: 7, 5, 6, 4, 12 New Drug: 3, 6, 4, 2, 1 Apply Mann-Whitney test, to check if there is a difference in the number of episodes of shortness of breath over a 1-week period in participants receiving the new drug as compared to those receiving the placebo? Use 1% level of significance.	*1*	*2*	**7**

			OR			
	6	a)	The following data represents the time, in minutes, that a patient has to wait during 12 visits to a doctor's office before being seen by the doctor: 17, 15, 20, 20, 32, 28, 12, 26, 25, 25, 35, 24. Apply Wilcoxon's signed test to test the doctor's claim that the median waiting time for her patient's is not more than 20 minutes before being admitted to the examination room. Use 5% level of significance.	*1*	*2*	**6**
		b)	A pharmaceutical company conducts an experiment to test the effect of a new cholesterol medication. The company selects 15 subjects randomly from a larger population. Each subject is randomly assigned to one of three treatment groups. Within each treatment group, subjects receive a different dose of the new medication. In Group 1, subjects receive 0 mg/day; in Group 2, 50 mg/day; and in Group 3, 100 mg/day. After 30 days, doctors measure the cholesterol level of each subject. The results for all 15 subjects appear in the table below: **Group 1,0 mg** 210 240 270 270 300 **Group 2,50 mg** 210 240 240 270 270 **Group 3, 100 mg** 180 210 210 210 240 Apply ANOVA to check whether the dosage level have a significant effect on cholesterol level? Use 1% level of significance.	*1*	*2*	**7**
		c)	One way to measure a person's fitness is by their body fat percentage. Average body fat percentages vary by age and the normal range for men is 15-20% body fat and that for women is 20-25% body fat. Our sample data is from a group of men and women who did workouts at a gym three times a week for a year. Then, their trainer measured the body fat. The table below shows the data. Men 13.3 6 20 8 14 19 18 25 16 24 15 1 15 Women 22 16 21.7 21 30 26 12 23.2 28 23 - - - At 5% level of significance, can we conclude that the mean body fat is the same for the larger population of men and women at the gym? Assuming the population variance are same.	*1*	*2*	**7**
			UNIT - 4			
	7	a)	Explain the following types of variables with an example. (i) Intervening variable (ii) Extraneous variable (iii) Composite variable	*1*	*1*	**6**
		b)	Briefly explain the advantages and disadvantages of the cohort design.	*1*	*1*	**7**

	c)	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. 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	*1*	*2*	**7**
		OR			
8	a)	Discuss the four levels of measurement scales. Give examples of each and explain in detail.	*1*	*1*	**6**
	b)	Researchers are conducting a prospective cohort study of the association between being an office worker who uses a computer daily and carpal tunnel syndrome. A total of 300 exposed and 300 unexposed participants are enrolled and followed for 10 years. A total of 25 exposed and 17 unexposed had the outcome of interest over the follow-up period. (i) What is the relative risk for developing carpal tunnel syndrome? (ii) What is the incidence attributable to daily computer use? (iii) If 60% of the population uses a computer daily at work, how much carpal tunnel could we prevent if we implemented a national work-place ergonomics program (and thus eliminated the exposure of daily computer use)?	*1*	*2*	**7**
	c)	In a school, the intelligent students are admitted to batch *A* of a class and the others are admitted to batch *B*. The students in the two batches are aged as follows. Compare their average age and variations. $Age(years)$: 10 | 11 | 12 | 13 | 14 | 15 $students(Batch\ A)$: 14 | 20 | 11 | 2 | 2 | 1 $students(Batch\ B)$: 2 | 12 | 11 | 13 | 4 | 2	*1*	*2*	**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 farms: 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. Blocks (Lands) | Fertilizers: A | B | C | D I: 10 | 12 | 14 | 13 II: 16 | 20 | 24 | 25 III: 18 | 21 | 26 | 28 Apply RBD to analyze data at 1% level of significance.	*1*	*2*	**10**

		b)	Apply Latin square design to estimate the missing value and hence construct the ANOVA table at 5% level of significance.	*1*	*2*	**10**

A-11	C-13	D-10	B-7	E-4
C-10	D-5	B-?	E-11	A-12
D-9	B-4	E-8	A-13	C-10
B-7	E-9	A-7	C-11	D-12
E-10	A-4	C-7	D-10	B-11

			OR			
	10	a)	An experiment was conducted on the yield of potatoes in a randomized block design with four replications. Analyze the following 2^2-factorial design at 5% level of significance.	*1*	*2*	**10**

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

		b)	A manufacturing firm wants to investigate the effects of 5 colour additives on the acting time of a new concrete mix. Variations in the setting times can be expected from day-to-day changes and humidity and also from the different workers who prepare the test molds. The data is given below with the letters A, B, C, D and E representing the five additives. The setting times, in hours, for the 25 molds are shown below.	*1*	*2*	**10**

Worker	Day				
	1	2	3	4	5
1	D-10.7	E-10.3	B-11.2	A-10.9	C-10.5
2	E-11.3	C-10.5	D-12	B-11.5	A-10.3
3	A-11.8	B-10.9	C-10.5	D-11.3	E-7.5
4	B-14.1	A-11.6	E-11	C-11.7	D-11.5
5	C-14.5	D-11.5	A-11.5	E-12.7	B-10.9

Apply Latin square design to analyze the data at 1% level of significance.

****