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: VI**
Branch: Institutional Elective **Duration: 3 hrs.**
Course Code: 23MA6OESFE **Max Marks: 100**
Course: MATHEMATICAL STATISTICS FOR ENGINEERS

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.

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 - 1

(*CO*, *PO*, **Marks**)

1 a) Discuss the four levels of measurement scales. Give examples of each and explain in detail. (*1*, *1*, **6**)

b) In order to find the average annual yield of coconuts per tree in a large coconut plantation, 1470 trees were randomly chosen and the yield were noted as below: (*1*, *1*, **7**)

No. of coconuts	120	130	140	150	160	170	180	190	200
No. of trees	15	78	273	369	341	207	133	43	11

Also, compute the co-efficient of mean deviation with respect to median.

c) The following are the distributions of lives of electric bulbs manufactured by two firms. Compare their mean and coefficient of variation. (*1*, *1*, **7**)

Life(hours)	800-899	900-999	1000-1099	1100-1199	1200-1299
No. of bulbs A	8	21	14	4	3
No. of bulbs B	16	24	40	16	4

OR

2 a) Treating the number of letters in each word in the following passage as the variable *x*, prepare the frequency distribution table: (*1*, *1*, **6**)

” *Morning dew clings to petals bright, Birdsong breaks the hush of night, Daylight spills its golden light*”.

Find the mean, median and mode of the distribution.

b) The following data relates to the number of apples per box in 100 boxes. (*1*, *1*, **7**)

No. of apples per box	120	121	122	123	124	125	126	127	128
No. of boxes	2	4	7	14	21	22	19	8	3

Calculate the co-efficient of quartile deviation.

c) The following data relates to the marks scored by students of a class in an examination. Compute the median and mode. (*1*, *1*, **7**)

% marks	<10	<20	<30	<40	<50	<60	<70
No. of students	4	16	20	65	85	97	100

			UNIT - 2			
	3	a)	Lots of 40 components each are called acceptable if they contain no more than 3 defectives. The procedure for sampling the lot is to select 5 components at random (without replacement) and to reject the lot if a defective is found. What is the probability that exactly one defective is found in the sample if there are 3 defectives in the entire lot? Also, find the expected value and the variance of the number of defectives in the sample.	*2*	*2*	**6**
		b)	The life (in hours) of a magnetic resonance imagining machine (MRI) is modelled by a Weibull distribution $f(t)=\alpha_1\beta t^{\beta-1}e^{-\alpha_1 t^{\beta}}$ with parameters $\alpha_1=(500)^{-2}$ and $\beta=2$ hours. (i) Determine the mean life of the MRI. (ii) Determine the variance of the life of the MRI. (iii) What is the probability that the MRI fails before 250 hours?	*2*	*2*	**7**
		c)	The plant manager does not know X, but from past experience she expects this probability to be equal to 33%. Furthermore, she quantifies her uncertainty about X by attaching a standard deviation of 5.5%. After consulting with an expert in statistics, the manager decides to use a Beta distribution to model her uncertainty about X. How should she set the two parameters of the distribution in order to match her priors about the expected value and the standard deviation of X ?	*2*	*2*	**7**
			OR			
	4	a)	Derive an expression for the mean and variance of a Uniform distribution of a discrete variable.	*2*	*1*	**6**
		b)	Suppose that it's winter in Toronto and we know that on average, it snows once every 10 days. Furthermore, suppose that 'it snowing' is an exponential distribution. (i) What is the expected time for the next four snow days to occur? (ii) What is the variance time for the next four snow days to occur? (iii) What is the probability that the next four snow days will occur next week?	*2*	*2*	**7**
		c)	Rate data often follow a lognormal distribution. Average power usage (dB per hour) for a particular company is studied and is known to have a lognormal distribution with parameters $\mu=4$ and $\sigma=2$. (i) What is the probability that the company uses more than 270dB during any particular hour? (ii) What is the mean and variance power usage?	*2*	*2*	**7**
			UNIT - 3			
	5	a)	A sample of three soccer players has to be chosen from a group of five players with numbers 23, 57, 12, 86, 39. Consider all possible samples without replacement and find (i) Population mean and standard deviation. (ii) Expected value and standard deviation of the sampling distribution of mean	*2*	*2*	**6**

		b)	Suppose that the population distribution of the gripping strengths of industrial workers is known to have a mean of 110 and standard deviation of 10. For a random sample of 75 workers, what is the probability that the sample mean gripping strength will be: (i) Between 109 and 112? (ii) Greater than 111?	*2*	*2*	**7**
		c)	The geometric random variable X has a probability distribution function $f(x)=(1-p)^{x-1}p,\ X=1,2,3,\ldots$. obtain the moment generating function and hence find the mean and variance.	*2*	*2*	**7**
			OR			
	6	a)	The lifespans of a type of LED bulb are normally distributed with a mean of 1200 hours and a standard deviation of 100 hours. A factory has 10,000 such bulbs in stock. If 400 random samples of size 64 are drawn: Determine the mean and standard deviation of the sampling distribution of sample means when sampling is done: (i) With replacement (ii) Without replacement (iii) What is the probability that the sample mean lifespan for one of these samples lies between 1180 hours and 1220 hours, assuming sampling is done with replacement?	*2*	*2*	**6**
		b)	Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50. (i) Apply Markovian inequality to find the probability that this week's production will exceed 75? (ii) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be between 40 and 60 by applying Chebyshev inequality?	*2*	*2*	**7**
		c)	A continuous random variable X has the following probability distribution $f(x)=3xe^{-5x},\ x>0$. obtain the moment generating function for X and hence find the mean and variance for X.	*2*	*1*	**7**
			UNIT - 4			
	7	a)	Suppose it is known that a sample consisting of the values 0.46, 0.38, 0.61, 0.82, 0.59, 0.53, 0.72, 0.44, 0.59, 0.60 comes from a population with the density function $f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta} & x>0\\ 0 & otherwise\end{cases}$. Find the estimator of θ by the method of moments.	*2*	*1*	**6**
		b)	Suppose $X_1, X_2, \ldots, X_n$ are independent $Uniform[0,\theta]$ random variables. Let $\hat{\theta}=\frac{3(X_1+X_2+\ldots+X_n)}{n}$ be an estimator of θ. Find the bias, variance and risk of $\hat{\theta}$.	*2*	*1*	**7**

c) The following data relate to the number of items produced per shift by two workers A and B for a number of days: 2 2 **7**

A	26	37	40	35	30	30	40	26	30	35	45
B	19	22	24	27	24	18	20	19	25	--	--

Assuming that the number of the items produced by both the workers follows normal distribution, estimate 95% confidence interval for population variance.

OR

8 a) A sample of 200 voters is chosen at random from all voters in a given city. 60% of them were in favor of a particular candidate. If large number of voters cast their votes, then find 99% confidence intervals for the proportion of voters in favor of a particular candidate. 2 2 **6**

b) The following are monthly family income in Rs. For a random sample of 16 families from a normal population:
381, 476, 372, 821, 507, 716, 524, 813, 578, 623, 739, 685, 830, 593, 364, 781. Obtain the maximum likelihood estimator of average and variance monthly income per family. 2 2 **7**

c) Suppose X and Y are independent $Exp\left(1/\theta\right)$ and $Uniform[0,\theta]$ random variables. Let $\hat{\theta} = aX + bY$ denote a class of estimators of θ where a and b are constants. 2 1 **7**

(i) Determine the bias and risk of $\hat{\theta}$.

(ii) Determine the condition involving a and b such that is unbiased for $\hat{\theta}$.

UNIT - 5

9 a) Suppose that 100 randomly selected families and 200 randomly selected singles were asked what type of car they drove: sport, sedan, hatchback, truck, van/SUV. The results are shown in table. Apply Chi-square test to test whether families and singles have the same distribution of cars? Test at a level of significance of 0.01. 2 2 **6**

	Sport	Sedan	Hatchback	Truck	Van/SUV
Family	5	15	35	17	28
Single	45	65	37	46	7

b) It is claimed that a new diet will reduce a person's weight, in a period of 2 weeks. The weights of 10 women who followed this diet were recorded before and after a 2-week period, yielding the following data: 2 2 **7**

Woman	1	2	3	4	5	6	7	8	9	10
Before	58.5	60.3	61.7	69.0	64.0	62.6	56.7	63.6	68.2	59.4
After	60.0	54.9	58.1	62.1	58.5	59.9	54.4	60.2	62.3	58.7

Apply Wilcoxon signed rank test to test the hypothesis that the diet reduces the median weight. Use 5% level of significance.

		c)	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. At 5% level of significance, is there any significant indication that supports the company's claim of the effectiveness of the medicine?	2	2	**7**
			OR			
	10	a)	The operations manager of a company that manufactures tires wants to determine whether there are any differences in the quality of workmanship among the three daily shifts. She randomly selects 496 tires and carefully inspects them. Each tire is either classified as perfect, satisfactory or defective, and the shift that produced it is also recorded. The two categorical variables of interest are: shift and condition of the tire produced. Do these data provide sufficient evidence at the 5% significance level to infer that there are differences in quality among the three shifts by applying Chi-square test?	2	2	**6**

	Perfect	**Satisfactory**	**Defective**
Shift 1	106	124	1
Shift 2	67	85	1
Shift 3	37	72	3

		b)	A fishing line is being manufactured by two processes. To determine if there is a difference in the mean breaking strength of the line, 10 pieces by each process are selected and then tested for breaking strength. The results are as follows:	2	2	**7**

Process 1	10.4	9.8	11.5	10.0	9.9	9.6	10.9	11.8	9.3	10.7
Process 2	8.7	11.2	9.8	10.1	10.8	9.5	11.0	9.8	10.5	9.9

At 0.01 level of significance, determine if there is a difference between the mean breaking strengths of the lines manufactured by the two processes by applying Mann -Whitney test.

		c)	Three nutrients were examined for the height of seedlings and the following data is recorded:	2	2	**7**

Height of seedlings

Nutrients A:	22	20	21	18	16	14
Nutrients B:	12	14	15	10	9	--
Nutrients C:	7	9	7	6	--	--

Perform one-way ANOVA at 1% level of significance.

******