U.S.N.										

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

Autonomous Institute Affiliated to VTU

October 2024 Supplementary Examinations

Programme: B.E. **Semester: IV**
Branch: Artificial Intelligence and Machine Learning **Duration: 3 hrs.**
Course Code: 24AM4PCIST **Max Marks: 100**
Course: Inferential Statistics

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)	Why investigators use samples to represent a population in statistical analysis and experimental studies? Discuss the reasons.	2	1	**5**
	b)	Given $X_1, X_2, \ldots, X_n$ be a random sample from population $N(\mu,1)$. Show that $t = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2$ is an unbiased estimator of $\mu^2 + 1$.	2	2	**7**
	c)	Given X_1, X_2 and X_3 to be a random sample of size 3 from a population with mean 'm' and variance 'k'. Given the estimators $T_1=X_1+X_2+bX_3$ and $T_2=2X_1 - 4X_2+aX_3$. i) Find the values of 'a' and 'b' such that T_1 and T_2 are unbiased for 'm' ii) Find the relative efficiency of T_1 with respect to T_2	2	2	**8**
		UNIT - II			
2	a)	Illustrate Stratified and Systematic sampling techniques used in Research study.	1	1	**6**
	b)	Consider a sample size of n=2 from a population of 5 consisting of 2, 3, 6, 8 and 11. Verify that Simple Random Sampling Without Replacement (SRSWOR) gives better estimate of population mean than Simple Random Sampling with Replacement (SRSWR).	1	1	**8**
	c)	The quality of ingredients used by a population of 20 Biscuit factories in strata are: 33, 39, 37, 31, 42, 68, 61, 43, 29, 31, 49, 41, 38, 40, 41, 38, 54, 32, 45, 42 i) Select all possible systematic sample of size 5. ii) Show that the sample mean is an unbiased estimator of population mean.	1	1	**6**
		UNIT - III			
3	a)	Write the procedure of Rejection Region Method to test Hypothesis of large samples.	3	1	**6**

		b)	Differentiate the following: i. Acceptance and Rejection region. ii. One-tailed and Two-tailed tests.	3	1	**7**
		c)	In a study to estimate the proportion of accidents in a major and minor cities due to construction, it was found that 158 of 400 major city residents and 155 of 500 minor city residents were in favor of construction. Find 95% confidence interval individually for proportion of major and minor city residents who favor the construction.	3	2	**7**
			UNIT - IV			
	4	a)	Describe and provide the z - test statistic used to determine the difference between two population means.	3	1	**5**
		b)	A sample of 900 members is found to have a mean of 3.4 cm. Can it be reasonably regarded as a sample from a large population of mean 3.25 cm and standard deviation, 2.61 cm? Solve and test at α=5% level of significance.	3	2	**7**
		c)	In two large population, there are 30% and 25% respectively of blue-eyed people. Is this difference likely to be hidden in the sample of 1200 and 900 respectively from the two population. Solve and test at α=5% level of significance.	3	2	**8**
			OR			
	5	a)	Discuss and provide the z - test statistic used to determine single population Proportion.	3	1	**5**
		b)	It is observed in a survey that 15% of households in a certain city indicated that they owned a washing machine. The survey based on a random sample of 900 households was taken and it was found that 189 households had a washing machine. Can we conclude that there has been a significant increase in the sale of washing machines at 5% level of significance?	3	2	**7**
		c)	The mean weight of 50 male students who showed above average participation in school athletics was 68.2 kgs with a standard deviation of 2.5 kgs. While 50 male students who showed no interest in such participation had a mean weight of 67.5 kgs with a standard deviation of 2.8 kgs. Test the hypothesis that male students who participate in school athletics are healthier than other students. Solve and test at α=5% level of significance.	3	2	**8**
			UNIT - V			
	6	a)	Define t-statistic and elaborate its procedure for testing equality of means for small sample.	3	1	**5**
		b)	The average breaking strength of the steel rods is specified to be 18.5 thousand kg. For this, a sample of 14 rods were tested. The mean and standard deviation obtained were 17.85 and 1.955, respectively. Solve and test the significance of deviation at α=5% level.	3	2	**7**

<table>
<tr><td></td><td></td><td>c)</td><td>The following figures relate to the number of units of an item produced per shift by two workers A and B for a number of days:
<table>
<tr><td>A</td><td>19</td><td>22</td><td>24</td><td>27</td><td>24</td><td>18</td><td>20</td><td>19</td><td>25</td><td></td><td></td></tr>
<tr><td>B</td><td>26</td><td>37</td><td>40</td><td>35</td><td>30</td><td>30</td><td>40</td><td>26</td><td>30</td><td>35</td><td>45</td></tr>
</table>
Can it be inferred that worker A is more stable compare to worker B? Solve and test at 5% level of significance with F-test.</td><td>3</td><td>2</td><td>8</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>7</td><td>a)</td><td>Describe the characteristics of the F distribution and provide its critical values for a two-tailed test. Include a neat, labeled diagram of its probability density function (pdf) indicating the critical values.</td><td>3</td><td>1</td><td>5</td></tr>
<tr><td></td><td></td><td>b)</td><td>Two types of batteries X and Y are tested for their length of life and the following result are obtained:
<table>
<tr><td>Battery</td><td>Sample size</td><td>Mean (hrs.)</td><td>Variance (hrs.)</td></tr>
<tr><td>X</td><td>10</td><td>1000</td><td>100</td></tr>
<tr><td>Y</td><td>12</td><td>2000</td><td>121</td></tr>
</table>
Is there a significant difference in the two means? Solve and test at 5% level of significance.</td><td>3</td><td>2</td><td>7</td></tr>
<tr><td></td><td></td><td>c)</td><td>70 accidents that occur during various days of the week as follows:
<table>
<tr><td>Day</td><td>Sun</td><td>Mon</td><td>Tue</td><td>Wed</td><td>Thu</td><td>Fri</td><td>Sat</td></tr>
<tr><td>Accidents</td><td>7</td><td>8</td><td>11</td><td>12</td><td>5</td><td>13</td><td>14</td></tr>
</table>
Solve and find whether the accidents are uniformly distributed over the week. Solve and test at 5% level of significance.</td><td>3</td><td>2</td><td>8</td></tr>
</table>

******