U.S.N.										

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

Autonomous Institute Affiliated to VTU

January / February 2025 Semester End Main Examinations

Programme: B.E.
Branch: INSTITUTION ELECTIVE
Course Code: 23MA6OESFE
Course: Mathematical Statistics for Engineers

Semester: VI
Duration: 3 hrs.
Max Marks: 100

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.

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.

<table>
<tr><th></th><th></th><th>UNIT - 1</th><th><i>CO</i></th><th><i>PO</i></th><th>Marks</th></tr>
<tr><td>1.</td><td>a)</td><td>Treating the number of letters in each word in the following passage as the variable x, prepare the frequency distribution table:
” The reliability of data must always be examined before any attempt is made to base conclusions upon them. This is true of all data, but particularly so of numerical data, which do not carry their quality written large on them. It is a waste of time to apply the refined theoretical method of statistics to data which are suspect from the beginning.”
Find the mean, median and mode of the distribution.</td><td><i>2</i></td><td><i>2</i></td><td>6</td></tr>
<tr><td></td><td>b)</td><td>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.
<table>
<tr><td>Wages</td><td>50-100</td><td>100-150</td><td>150-200</td><td>200-250</td><td>250-300</td><td>300-350</td></tr>
<tr><td>Shift A</td><td>3</td><td>8</td><td>24</td><td>63</td><td>102</td><td>50</td></tr>
<tr><td>Shift B</td><td>32</td><td>44</td><td>62</td><td>41</td><td>37</td><td>34</td></tr>
</table></td><td><i>2</i></td><td><i>2</i></td><td>7</td></tr>
<tr><td></td><td>c)</td><td>Explain the different types of scales of measurement used in the descriptive statistics.</td><td><i>1</i></td><td><i>1</i></td><td>7</td></tr>
<tr><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>a)</td><td>An incomplete frequency distribution is given as follows:
<table>
<tr><td>Variables</td><td>Frequency</td><td>Variables</td><td>Frequency</td></tr>
<tr><td>10-20</td><td>12</td><td>50-60</td><td><i>45</i></td></tr>
<tr><td>20-30</td><td>X</td><td>60-70</td><td>Y</td></tr>
<tr><td>30-40</td><td><i>34</i></td><td>70-80</td><td>18</td></tr>
<tr><td>40-50</td><td>65</td><td>Total</td><td>229</td></tr>
</table>
Given that the median value is 46, determine the value of <i>X</i> and <i>Y</i>.</td><td><i>2</i></td><td><i>1</i></td><td>6</td></tr>
</table>

B.M.S.C.E. - ODD SEM 2024-25

<table>
<tr><td></td><td></td><td>b)</td><td>The median and mode of the following wages are known to be Rs. 33.5 and Rs. 34 respectively. Determine the value of x, y and z, given total frequency is 230.
<table>
<tr><td>Class</td><td>0-10</td><td>10-20</td><td>20-30</td><td>30-40</td><td>40-50</td><td>50-60</td><td>60-70</td></tr>
<tr><td>Frequency</td><td>4</td><td>16</td><td>x</td><td>y</td><td>z</td><td>6</td><td>4</td></tr>
</table></td><td>2</td><td>1</td><td>7</td></tr>
<tr><td></td><td></td><td>c)</td><td>A survey was conducted on the inmates of a prison and the following data were found:
<table>
<tr><td><i>Period of jail (in months)</i></td><td>6</td><td>10</td><td>12</td><td>20</td><td>24</td><td>36</td><td>40</td><td>48</td></tr>
<tr><td><i>No. of prisoners</i></td><td>7</td><td>11</td><td>14</td><td>36</td><td>16</td><td>9</td><td>5</td><td>1</td></tr>
</table>
Calculate the coefficient of quartile deviation and mean deviation from median.</td><td>2</td><td>1</td><td>7</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - 2</td><td></td><td></td><td></td></tr>
<tr><td></td><td>3</td><td>a)</td><td>A representative from the National Football League's Marketing Division randomly selects people on a random street in Kansas City, Kansas until he finds a person who attended the last home football game. Let the probability that he succeeds in finding such a person is 0.20. And, let X denote the number of people he selects until he finds his first success.
i) What is the probability that the marketing representative must select 4 people before he finds one who attended the last home football game?
ii) What is the probability that the marketing representative must select more than 6 people before he finds one who attended the last home football game?
iii) How many people should we expect (that is, what is the average number) the marketing representative needs to select before he finds one who attended the last home football game?</td><td>2</td><td>2</td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>A bus travels between two cities A and B which are 100 miles apart. If the bus has a breakdown, the distance X of the point of breakdown from the city A has a uniform distribution $U[0,100]$. There are service garages in the city A, city B and midway between the two cities such that in case of a breakdown a tow truck is sent from the garage nearest to the point of breakdown.
(i) What is the probability that the tow truck has to travel more than 10 miles to reach the bus?
(ii) Would it be more “efficient” if the three service garages were placed at 25, 50 and 75 miles from city A, apart from service garages at city A and city B?</td><td>2</td><td>2</td><td>7</td></tr>
<tr><td></td><td></td><td>c)</td><td>A variate t is said to be a Students’ t- distribution on n degrees of freedom if its probability density function is given by $f(t)=\frac{k}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}, -\infty < t < \infty$. Find the value of k and show that $f(t)$ is a probability density function of t-distribution.</td><td>1</td><td>1</td><td>7</td></tr>
</table>

B.M.S.C.E. - ODD SEM 2024-25

			OR			
	4	a)	A small voting district has 101 female voters and 95 male voters. A random sample of 10 voters is drawn. i) What is the probability exactly 7 of the voters will be female? ii) What is the probability that at least one of the voters will be female? iii) Find the mean and variance for the given sample.	*2*	*2*	**6**
		b)	The lifetime X (in months) of a computer has a gamma distribution with mean 24 months and standard deviation 12 months. Find the probability that the computer will (i) last between 12 and 24 months (ii) last at most 24 months.	*2*	*2*	**7**
		c)	Derive an expression for the mean and variance of Chi Square distribution.	*1*	*1*	**7**
			UNIT - 3			
	5	a)	A rectangular field is divided into squares of side 1 meter. At one time of the year the number of snails in the field can be modeled by a Poisson distribution with mean $2.25 snail/m^2$. A random sample of 100 squares is observed and the number of snails in each square is counted. Find the probability that the sample mean number of snails is at most 2.5.	*2*	*2*	**6**
		b)	Let X be the number of screws delivered to a box by an automatic filling device with mean 25 and variance 16. There are problems with too many (giving away free product) or too few (potential irritated customer) screws in a box. Apply Chebyshev's inequality to find: i) a lower bound for $P(17 < x < 33)$. ii) an upper bound for $P(\|x-25\| \geq 12)$.	*2*	*2*	**7**
		c)	Derive the moment generating function of an exponential distribution and hence obtain its mean and variance.	*1*	*1*	**7**
			OR			
	6	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. Construct the sampling distribution for this process. Also, find the population mean and standard error of the sampling distribution of sample scores obtained.	*1*	*1*	**6**

B.M.S.C.E. - ODD SEM 2024-25

<table>
<tr><td></td><td></td><td>b)</td><td>Suppose that it is known that the amount of uranium dug in a uranium mine during one week is a random variable with mean 20 kg.
(i) Apply Markov’s inequality to find a bound on the probability that the week’s production will exceed 24 kg?
(ii) If the Variance of the amount of uranium dug in the mine equals 16 kg. Apply Chebyshev’s inequality to find a lower bound on the probability that week’s production will be between 8 and 32 kg?</td><td>1</td><td>1</td><td>7</td></tr>
<tr><td></td><td></td><td>c)</td><td>A continuous random variable X has the following probability distribution $f(x) = 3xe^{-5x},\ x > 0$.
(i) Find the moment generating function for X.
(ii) Find the mean and variance for X.</td><td>1</td><td>1</td><td>7</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - 4</td><td></td><td></td><td></td></tr>
<tr><td></td><td>7</td><td>a)</td><td>The sample values from the population with probability function $f(x) = (1+\theta)x^{\theta}, 0 < x < 1, \theta > 0$ are given as 0.46, 0.38, 0.62, 0.82, 0.59, 0.53, 0.72, 0.44, 0.54. Obtain an estimate for θ by method of moments.</td><td>1</td><td>1</td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>A study is to be made of the relative effectiveness of two kinds of cough medicines in increasing sleep. Six people with colds are given medicine A the first night and medicine B the second night. Theirs of sleep each night are recorded. The data are:
<table>
<tr><td>Medicine A</td><td>4.8</td><td>4.1</td><td>5.8</td><td>4.9</td><td>5.3</td><td>7.4</td></tr>
<tr><td>Medicine B</td><td>3.9</td><td>4.2</td><td>5.0</td><td>5.9</td><td>5.4</td><td>7.1</td></tr>
</table>
Establish a 95% confidence interval for the mean increase in hours of sleep from medicine B to medicine A.</td><td>2</td><td>2</td><td>7</td></tr>
<tr><td></td><td></td><td>c)</td><td>Suppose $X_1, X_2, ...X_n$ are identically, independent Uniform random variable $[0,\theta]$. Let $\hat{\theta} = \frac{3(X_1 + X_2 + ... + X_n)}{n}$ be an estimator of θ. Find the bias, variance and risk of $\hat{\theta}$.</td><td>1</td><td>1</td><td>7</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>8</td><td>a)</td><td>Suppose 10 rats are used in a biomedical study where the rats are injected with cancer drug that is designed to increase their survival rate, the survival times, in months are 14, 17, 27, 18, 12, 8, 22, 13, 19 and 12. Assume the survival rate follows exponential distribution. Find the maximum likelihood estimate of the parameter.</td><td>1</td><td>1</td><td>6</td></tr>
</table>

B.M.S.C.E. - ODD SEM 2024-25

<table>
<tr><td></td><td></td><td>b)</td><td>Let $X_1, X_2, ..., X_n$ are independent $Bernoulli(p)$ random variables. Let $\hat{p} = \frac{X_1 + X_2 + ... + X_n + \sqrt{n}/2}{n + \sqrt{n}}$ be an estimator of θ. Find the bias, variance and risk of $\hat{p}$.</td><td>2</td><td>1</td><td>7</td></tr>
<tr><td></td><td></td><td>c)</td><td>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 $\bar{x}_1 = 30.68$ fluid ounces. Assume normality. Construct a 90% confidence interval on the mean difference in fill volume.</td><td>2</td><td>2</td><td>7</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - 5</td><td></td><td></td><td></td></tr>
<tr><td></td><td>9</td><td>a)</td><td>A study is run to evaluate the effectiveness of an exercise program in reducing systolic blood pressure in patients with pre-hypertension. A total of 10 patients with pre-hypertension enrol in the study, and their systolic blood pressures are measured. Each patient then participates in an exercise training program where they learn proper techniques and execution of a series of exercises. Patients are instructed to do the exercise program 3 times per week for 6 weeks. After 6 weeks, systolic blood pressures are again measured. The data are shown below.
<table>
<tr><td>Before</td><td>125</td><td>132</td><td>138</td><td>120</td><td>125</td><td>127</td><td>136</td><td>139</td><td>131</td><td>132</td></tr>
<tr><td>After</td><td>118</td><td>134</td><td>130</td><td>124</td><td>105</td><td>130</td><td>130</td><td>132</td><td>123</td><td>128</td></tr>
</table>
Apply Wilcoxon Signed rank test to determine if there is a difference in systolic blood pressures after participating in the exercise program as compared to before?</td><td>2</td><td>2</td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>The following figures relate to production in kg. of three varieties A, B and C of wheat sown in 12 plots.
<table>
<tr><td>A</td><td>14</td><td>16</td><td>18</td><td>-</td><td>-</td></tr>
<tr><td>B</td><td>14</td><td>13</td><td>15</td><td>22</td><td>-</td></tr>
<tr><td>C</td><td>18</td><td>16</td><td>19</td><td>19</td><td>2</td></tr>
</table>
Apply One-way ANOVA to test if there is any significant difference in the production of three varieties of wheat at 5% level of significance.</td><td>2</td><td>2</td><td>7</td></tr>
</table>

B.M.S.C.E. - ODD SEM 2024-25

<table>
<tr><td></td><td></td><td>c)</td><td>In a cross between rust-resistant and rust-susceptible varieties of oats, the F_3 families were compared for rust reaction in the seedling stage and in the field under ordinary epidemic conditions. The data are as follows:
<table>
<tr><th rowspan="2">Field Reaction</th><th colspan="3">Seedling Reaction</th></tr>
<tr><th>Resisting</th><th>Segregating</th><th>Susceptible</th></tr>
<tr><td>Resisting</td><td>142</td><td>51</td><td>7</td></tr>
<tr><td>Segregating</td><td>13</td><td>404</td><td>5</td></tr>
<tr><td>Susceptible</td><td>5</td><td>17</td><td>176</td></tr>
</table>
Apply Chi-square independence of attributes to test whether the rust reaction is independent in two stages at 1% level of significance.</td><td>2</td><td>2</td><td>7</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>At a certain date in a large city, 16 out of a random sample of 500 men were found to be drinkers. After the heavy increase in tax on intoxicants, another random sample of 100 men in the same city included 3 drinkers. Was the observed decrease in the proportion of drinkers significant? Use 5% level of significance.</td><td>2</td><td>2</td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>From the adult population of 4 large cities, random samples were selected and the number of married and unmarried men were recorded.
<table>
<tr><th></th><th>A</th><th>B</th><th>C</th><th>D</th></tr>
<tr><td>Married</td><td>137</td><td>164</td><td>152</td><td>147</td></tr>
<tr><td>Single</td><td>32</td><td>57</td><td>56</td><td>35</td></tr>
</table>
Apply Chi-square test for homogeneity to check whether there is a significant variation among the cities in the tendency of men marrying? Use 5% level of significance.</td><td>2</td><td>2</td><td>7</td></tr>
<tr><td></td><td></td><td>b)</td><td>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.
<table>
<tr><td>Placebo</td><td>7</td><td>5</td><td>6</td><td>4</td><td>12</td></tr>
<tr><td>New Drug</td><td>3</td><td>6</td><td>4</td><td>2</td><td>1</td></tr>
</table>
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.</td><td>2</td><td>2</td><td>7</td></tr>
</table>

******

B.M.S.C.E. - ODD SEM 2024-25