	GEETHANJALIINSTITUTEOFSCIENCE&TECHNOLOGY (ANAUTONOMOUSINSTITUTION) (ApprovedbyAICTE,NewDelhi&AffiliatedtoJNTUA,Ananthapuramu) (Accredited by NAAC with "A" Grade, NBA (EEE,ECE &ME) & ISO9001:2008CertifiedInstitution)

QUESTIONBANK(DESCRIPTIVE)

SubjectName:STATISTICAL METHODS FOR DATA SCIENCE

Course & Branch: B.Tech & CSE (DS)
Year&Semester: II B.TechII Semester
Regulation:RG23

UNIT-I

S.No	Question	[BTLevel][CO][Marks]
	2MarksQuestions (Short)	
1.	What is the formula of Mean and Variance of a discrete random variable?	L1, CO1, 2M
2.	Write formula to calculate mean and variance of a continuous random variable.	L1, CO1 , 2M
3.	If X is a continuous random variable and Y=a X +b, Prove that E(Y)=a E(X)+b.	L1 ,CO1, 2M
4.	The mean and variance of a binomial distribution are 4 and 4/3 respectively. Find n, p, q	L2, CO1, 2M
5.	write conditions of Poisson distribution	L1, CO1, 2M
6	Write probability of density function normal distribution	L1, CO1, 2M
7	Write the characteristics of a good estimator	L1, CO1, 2M
8	Define unbiasedness	L1, CO1, 2M
9	Define sufficient	L1, CO1, 2M
10	Define efficient estimator	
	Descriptive Questions(Long)	
1.	Two dice are thrown. Let X assign to each point (a, b) is S the maximum of its numbers I .e, X (a ,b)=max(a, b). Find the probability distribution. X is a random variable with X(s) = {1,2 ,3,4,5,6}. Also, find the mean and variance of the distribution	L2, CO1,10M
2.	A random variable X has the following probability function X=x 1 2 3 4 5 6 7 8 P(X=x) K 2K 3 K 4K 5K 6K 7K 8K Find the value of k and (i) Mean (ii) Variance (iii) P(2≤x≤5)	L2, CO1,10M
3.	If the probability density of a random variable is given by f(x) = kx^2e^{-x} for $x \geq 0$ 0 otherwise i)Find the value of k ii) mean and variance of the variable	L2, CO1,10M
4.	If the probability density of a random variable is given by f(x)= $[k\ (1-x^2),\ for\ 0 < x < 1]$, = 0,otherwise Find the value of k and the probabilities that a random variable having this probability density will take on a value (i) between 0.1 and 0.2 (ii) greater than 0.5.	L2, CO1,10M
5.	Out of 800 families with 4 children each, how many families would be expected to have (a) 2 boys and 2 girls (b) at least one boy (c) no girl (d) at most 2 girls? Assume equal probabilities for boys and girls	L2, CO2,10M
6.	a)As only 3 students came to attend the class today, find the probability for exactly 4 students to attend the classes tomorrow. b)If x is Poisson varieties such that p(x = 0) = p (x=1) and using recurrence relation formula fimd the probabilities at x= 1,2,3,4,5	L2, CO1,10M

7	a) If x is normal vitiate with mean 30 and standard deviation with 5 . find the probabilities (i) $26 \leq x \leq 40$ (ii) $x \geq 45$ b) A sales tax has reported that the average sales of the 500 business that has deal during with a year 36,000 with standard deviation 10,000 Assuming that the sales in these distributed in normal distribution Find (i) The number of business as sales which RS 40,000 (ii)the percentage of business sales of which are likely range between RS 30,000 and RS 40, 000	
8	In a Normal distribution, 7% of the items are under 35 and 89% are under 63. Determine the mean and variance of the distribution	
9	Show that the sample variance of a sample drawn from population given by the expression the s^2= 1/n $\sum(x_i - \mathrm{x})$ is an unbiased estimator of the population variance	
10	Let $x_1, x_2, x_3, \ldots.. x_n$ is a random sample taken from a population with probability density function $f(x) = \theta\, x^{\theta-1}$, $0 < x < 1, \theta > 0$.Find a sufficient estimator for θ.	

UNIT -III

<table>
<tr><th>S.No.</th><th>Question</th><th>[BTLevel][CO][Marks]</th></tr>
<tr><td colspan="3">2MarksQuestions (Short)</td></tr>
<tr><td>1.</td><td>Define estimation, estimate and estimator</td><td>L1,CO2,2M</td></tr>
<tr><td>2.</td><td>List the properties of the Maximum Likelihood Estimator.</td><td>L1,CO2,2M</td></tr>
<tr><td>3.</td><td>Define normal equations for two-variable relationships Y on X1, X2 in the method of least squares</td><td>L1,CO2,2M</td></tr>
<tr><td>4.</td><td>Define the Asymptotic Maximum Likelihood Estimator.</td><td>L1,CO2,2M</td></tr>
<tr><td>5.</td><td>What is the formula for the modified minimum chi-square method?</td><td>L1,CO2,2M</td></tr>
<tr><td colspan="3">DescriptiveQuestions(Long)</td></tr>
<tr><td>11.</td><td>Describe the method of maximum likelihood estimation.</td><td>L2,CO2, 10M</td></tr>
<tr><td>12.</td><td>In a watch repair shop, the service time in minutes is 14,17,27,18,12,8,22,13,19 and 12. Give a maximum likelihood estimate of mean service time with the assumption that the service time follows an exponential distribution with parameter λ.</td><td>L2,CO2, 10M</td></tr>
<tr><td>13.</td><td>A simple random sample of size n is taken from the probability density function
⟦f(x)= 2θx$e^{-\theta x^2}$,x>0,θ>0 is an unknown parameter.
Calculate the estimator of θ by the method of moments.</td><td>L2, CO2,10M</td></tr>
<tr><td>14.</td><td>Explain the method of moment's estimation.</td><td>L2, CO2,10M</td></tr>
<tr><td>15</td><td>Let $x_1, x_2, x_3 ... x_n$ be a random sample from the discrete distribution
$P(X_1=1)=(2(1-\theta))/(2-\theta)$, $P(X_2=2)=\theta/(2-\theta)$, where θ ϵ(0,1) is unknown.
Find the estimator θ by the method of moments</td><td>L2, CO2,10M</td></tr>
<tr><td>16</td><td>A training data set of 9 different values for mid semester (say x) and end semester (say y) values are given by
<table>
<tr><td>X</td><td>10</td><td>7</td><td>3</td><td>16</td><td>9</td><td>11</td><td>7</td><td>10</td><td>8</td></tr>
<tr><td>Y</td><td>42</td><td>39</td><td>32</td><td>50</td><td>44</td><td>55</td><td>43</td><td>37</td><td>43</td></tr>
</table>
Assuming a linear relationship y o x . Estimate the parameters by the method of least squares</td><td>L2, CO2,10M</td></tr>
<tr><td>17.</td><td>Past experiences shown the following result of productivity per hector with respective use of fertilizers and seeds. Assuming the linear relationship of y on x_1,x_2. Estimate the parameter from the given data
<table>
<tr><td>Fertilizer x_1</td><td>45</td><td>30</td><td>70</td><td>75</td><td>65</td><td>80</td></tr>
<tr><td>Seeds x_2</td><td>2</td><td>18</td><td>3</td><td>2.5</td><td>2</td><td>3</td></tr>
<tr><td>Productivity y</td><td>2000</td><td>2100</td><td>1800</td><td>1900</td><td>2400</td><td>2500</td></tr>
</table></td><td>L2, CO2,10M</td></tr>
<tr><td>18.</td><td>Consider a random sample of size n from an exponential distribution with parameter θ having p.d.f ⟦f(x,θ)=θe⟧^(-θx),x>0. Calculate the modified minimum chi-square estimator of θ based on the partitions A_1={x,0≤x≤0.5} , A_2={x,x>0.5}</td><td>L2, CO2,10M</td></tr>
<tr><td>19.</td><td>Suppose that the random sample has Normal distribution N(μ, σ2)
Find the maximum likelihood estimator (i) for μ when $\sigma 2 = 1$, (ii) for $\sigma 2$ when $\mu = 0$</td><td></td></tr>
<tr><td>20.</td><td>Suppose that the random sample has exponential distribution with parameters (α,β) f(x, α,β) $= y_0 e^{-\beta(x-\alpha)}$ $\alpha \le x < \infty, \beta > 0$, y_o be a constant . (i)find the constant y_o (ii) find the MLE for paramaters</td><td></td></tr>
</table>