Reg. No.															

Question Paper Code	12730

B.E. / B.Tech. - DEGREE EXAMINATIONS, APRIL / MAY 2024

Second Semester

Computer Science and Business Systems

20BSMA203 - STATISTICAL METHODS WITH LABORATORY

Regulations - 2020

Duration: 3 Hours | Max. Marks: 100

PART - A (10 × 2 = 20 Marks)

Answer ALL Questions

		Marks	K – Level	CO
1.	What is the line of regression of Y on X and Angle between two lines of regression?	2	K1	CO1
2.	What are the basic principles of the design of experiments?	2	K1	CO1
3.	Define type I and type II errors.	2	K1	CO2
4.	What are the applications of F-test?	2	K1	CO2
5.	What are the characteristics of estimators?	2	K1	CO3
6.	A random sample of 10 cadets of a centre is selected and measures their weights (in kg) which are given below: 48, 50, 62, 75, 80, 60, 70, 56, 52, 78 Determine an unbiased estimate of the average weight of cadets of the centre.	2	K2	CO3
7.	Write two assumptions made in non-parametric test.	2	K2	CO4
8.	Give the formula to find the rank correlation coefficient in case of tie rank.	2	K2	CO4
9.	When a time-series does is said to be a strictly stationary?	2	K1	CO5
10.	How does ARIMA forecasting work?	2	K2	CO5

PART - B (5 ×16 = 80 Marks)

Answer ALL Questions

11 a) Calculate the coefficient of correlation and obtain the lines of regression for the following data. *16 K3 CO1*

X	1	2	3	4	5	6	7	8	9
Y	9	8	10	12	11	13	14	16	15

Obtain an estimate of Y which should corresponding to the value =6.2.

OR

b)

	Machine Type				
		A	**B**	**C**	**D**
Workers	1	44	38	47	36
	2	46	40	52	43
	3	34	36	44	32
	4	43	38	46	33
	5	38	42	49	39

16 K3 CO1

The above data represent the number of units of production per day

turned out by different workers using 4 different types of machines. Test whether the five men differ with respect to mean productivity and test whether the mean productivity is the same for the four different machine types.

12. a) i) A sample of 900 members has a mean 3.4 cm and standard deviation 2.61 cm. Is the sample from a large population of mean 3.25 cms and standard deviation of 2.61 cms. Test at 5% level of significance. The value of Z at 5% level of significance is $|Z_\alpha| < 1.96$ *8 K3 CO2*

ii) A sample of 10 boys had the following IQ's: 70, 120, 110, 101, 88, 83, 95, 98, 100 and 107. Test whether the population IQ may be 100. *8 K3 CO2*

OR

b) i) The time taken by workers in performing a job by method I and method II is given below: *8 K3 CO2*

Method I	20	16	26	27	23	22	
Method II	27	33	42	35	32	34	38

Do the data show that the variances of time distribution from population from which these samples are drawn do not differ significantly?

ii) A manufacturer claimed that at least 95% of the equipment which he supplied to a factory conformed to specifications. An examination of a sample of 200 pieces of equipments revealed that 18 were faulty. Test his claim at a significance level 5% *8 K3 CO2*

13. a) i) Prove that $s^2 = \frac{\sum(X_i - \bar{X})^2}{n}$ is not an unbiased estimator of population variance σ^2. From that prove $S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1}$ is an unbiased estimator. *10 K3 CO3*

ii) Find the MLE of θ in $f(x) = \frac{1}{\theta}, 0 \le x \le \theta, \theta > 0$. *6 K3 CO3*

OR

b) For random sampling from normal population N (μ, σ^2), find the maximum likelihood estimators for μ and σ^2. *16 K3 CO3*

14. a) Independent random sample of 10 day scholar and 10 hostel students in an university is considered. Use Mann-Whitney test to determine if there is significant difference in age distribution of two samples. Use $\alpha = 0.05$ and test for any significant difference in age distribution of the two populations. *16 K3 CO4*

Day scholar	26	18	25	27	19	30	34	21	23	31
Hostellers	32	24	23	30	40	41	42	39	45	35

OR

b) i) In industrial production items are periodically inspected for defectives. The following is sequence of defective D and non-defective N items produced by these production lines DDNNNDNNDDNNNNNDDDNNDNNNNDND. Test whether defectives are occurring at random or not at 5% LOS. *8 K3 CO4*

ii) A sales manager collects the following salary statistics on this field sales force earning. He has both observed frequencies and expected frequencies. If the distribution of salaries is normal at 0.05 LOS can we conclude that the distribution of sales force earning is normal use Kolmogorov-Smirov test *8 K3 CO4*

Interval	25-30	31-36	37-42	43-48	49-54	55-60	61-65
Observed frequency	9	22	25	30	21	12	6
Expected frequency	6	17	32	35	18	13	4

15. a) For each of the following models: *16 K3 CO5*

(a)$X_t = 0.3X_{t-1} + Z_t$

(b)$X_t = Z_t - 1.3Z_{t-1} + 0.4Z_{t-2}$

(c)$X_t = 0.5X_{t-1} + Z_t - 1.3Z_{t-1} + 0.4Z_{t-2}$

Express the model using B notation and determine whether the model is stationary and/or invertible.

OR

b) i) Show that the auto correlation function of the second-order MA process *8 K3 CO5*

$$X_t = Z_t + 0.7Z_{t-1} - 0.2Z_{t-2}$$

is given by

$$\rho(k) = \begin{cases} 1 & k = 0 \\ 0.37 & k = \pm 1 \\ -0.13 & k = \pm 2 \\ 0 & \text{otherwise} \end{cases}$$

ii) For the model $(1 - 0.2B)(1 - B)X_t = (1 - 0.5B)Z_t$: *8 K3 CO5*

(a) Classify the model as an ARIMA (p, d, q) process (i.e. find p, d, q).

(b) Determine whether the process is stationary and invertible