NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: II - THEORY EXAMINATION (2021 - 2022)

Subject: Statistical Methods

Time: 3 Hours   Max. Marks: 100

General Instructions:

1. The question paper comprises three sections, A, B, and C. You are expected to answer them as directed.

2. Section A - Question No- 1 is 1 marker & Question No- 2 carries 2 mark each.

3. Section B - Question No-3 is based on external choice carrying 6 marks each.

4. Section C - Questions No. 4-8 are within unit choice questions carrying 10 marks each.

5. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.

SECTION A   20

1. Attempt all parts:-

1-a.   The difference between a statistic and the parameter is called :   (CO1)   1

                (a) Non -random

                (b) Probability

                (c) Sampling error

                (d) Random

1-b.   Standard deviation of sampling distribution of a statistic is called :   (CO1)   1

                (a) Serious error

                (b) Dispersion

                (c) Standard error

                (d) Difference

1-c.   (ANOVA) Analysis of variance is a statistical method of comparing the_____ of several   1
populations. (CO2)

                (a) Means

                (b) Variances

                (c) Standard Deviations

(d) None of the above

1-d. Which of the following statement is true ? (CO2) 1

(a) The correlation coefficient and the two regression coefficients have the same sign.

(b) The correlation coefficient and a regression coefficient have the same sign.

(c) The correlation coefficient and the two regression coefficients have the opposite sign.

(d) None of these

1-e. Bias of an estimator can be: (CO3) 1

(a) Negative

(b) Positive

(c) Zero

(d) Both A or B

1-f. The "Likelihood Equations" are: (CO3) 1

(a) The same as the "normal equations" associated with least squares estimation of the multiple linear regression model

(b) Guaranteed to have a unique solution if the sample data are independent

(c) Obtained by getting the second derivatives of the log-likelihood function with respect to each of the parameters, and setting these equal to zero

(d) The first-order conditions that we have to solve in order to maximize the likelihood function

1-g. A failing student is passed by an examiner, it is an example of: (CO4) 1

(a) Type-I error

(b) Type –II error

(c) Unbiased decision

(d) None of the above

1-h. The test which enables to compare the two independent samples by testing a hypothesis on the two population medians is: (CO4) 1

(a) Mann Whitney test

(b) T-test

(c) Z-test

(d) None of these

1-i. The additive model of the time series with the components T, S, C and I is (CO5) 1

(a) $y = T + S + C \times I$

(b) y=T+S×C×I

(c) y=T+S+C+I

(d) y=T+S×C+I

1-j.   The component of a time series attached to long term variation is trended as (CO5)   1

(a) Cyclic variation

(b) Secular variations

(c) Irregular variation

(d) Seasonal variations

2. Attempt all parts:-

2.a.   Define Standard Error ? (CO1)   2

2.b.   Explain two-way analysis of variance with a focus on its merits and demerits? (CO2)   2

2.c.   What is confidence interval? (CO3)   2

2.d.   What should be the size of samples on which Z-test is applied? (CO4)   2

2.e.   Explain forecasting briefly.   (CO5)   2

SECTION B   30

3. Answer any <u>five</u> of the following:-

3-a.   What are the advantages and disadvantages of simple random sampling? CO1   6

3-b.   How to perform Stratified Random Sampling? CO1   6

3-c.   Compute the Rank correlation coefficient for the following data: (CO2)   6

Person	A	B	C	D	E	F	G	H	I	J
Rank in Mathematics	9	10	6	5	7	2	4	8	1	3
Rank in Physics	1	2	3	4	5	6	7	8	9	10

3-d.   A driver keeps a record of the distance travelled and the amount of fuel in his tank on a long   6
journey. Draw the scatter graph for this data. (CO2)

Distance

Travelled   0   50   100   150   200   250   300
(km)

Fuel in Tank (litres)	80	73	67	61	52	46	37

3.e.   Differentiate between point estimate and interval estimate. (CO3)   6

3.f.   An automatic machine was designed to pack exactly 2.0 kg of oil. A sample of 100 tins was examined to test the machine. The average weight was found to be 1.94 kg with standard deviation 0.10 kg. Is the machine working properly? (CO4)   6

3.g.   What does the p, d, and q in ARIMA models means? (Explain briefly) (CO5)   6

<div align="center">

SECTION C   50

</div>

4. Answer any <u>one</u> of the following:-

4-a.   From the data given below   10

X	6	2	10	4	8
y	9	11	5	8	7

Find standard error of estimate (S). (CO1)

4-b.   Define sampling and various types of sampling with example. (CO1)   10

5. Answer any <u>one</u> of the following:-

5-a.   Obtain a regression plane by using multiple linear regression to fit the data given below: (CO2) ?   10

x	1	2	3	4
y	0	1	2	3
z	12	18	24	30

5-b.   Three variety A, B and C of wheat are grown in 5 plots each and the following yields per acre were obtained:. (CO2)   10

A	B	C
8	7	12
10	5	9
7	10	13
14	9	12
11	9	14

Set up a table of (ANOVA) analysis of variance and find out whether there is a significant difference between the mean yields of these varieties.(Test at 5% level of significance 3.88)

6. Answer any <u>one</u> of the following:-

6-a.   10

1. A random sample of n = 6 has the element 7, 11, 12, 13, 18 and 20. Compute a point estimate of

    i. Population mean

    ii. The population standard deviation

    iii. The standard error of the mean   (CO3)

6-b.   Explain sufficient statistics. When the sufficiency is called complete sufficiency. What is the   10
application of sufficiency in estimation? (CO3)

7. Answer any <u>one</u> of the following:-

7-a.   A large hospital hires most of its doctors from the two major universities. Over the last year,   10
hospital has been conducting test for the newly recruited doctors to determine which school
educate better. Based on the following scores, help the human resource department of the
hospital to decide whether the universities differ in quality. (Use Mann-Whitney U-Test,
$\alpha = 0.10$, critical value = 1.65). (CO4)

University A	99	83	89	64	98	85	61	79	91	87	88	
University B	96	90	97	94	86	95	68	78	93	56	76	84

7-b.   Phillips company claims that the length of life of its electric bulb is 2000 hours with standard   10
deviation of 30 hours. A random sample of 25 showed an average life of 1940 hours with a
standard deviation of 25 hours. At 5% level of significance can we conclude that the sample
has come from a population with mean of 2000 hours? (CO4)

8. Answer any <u>one</u> of the following:-

8-a.   Explain ARIMA briefly? (CO5)   10

8-b.   Explain the type of time series additive and multiplicative models? Also explain their   10
components? (CO5)