

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: II - CARRY OVER THEORY EXAMINATION - AUGUST 2023

Subject: Statistical Methods

Time: 3 Hours

Max. Marks: 100

General Instructions:

IMP: Verify that you have received the question paper with the correct course, code, branch etc.

1. This Question paper comprises of **three Sections -A, B, & C**. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.
2. Maximum marks for each question are indicated on right -hand side of each question.
3. Illustrate your answers with neat sketches wherever necessary.
4. Assume suitable data if necessary.
5. Preferably, write the answers in sequential order.
6. 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. When a random starting point is chosen, followed by every nth individual, the sampling method is **(CO1)** 1
- (a) simple random sampling
 - (b) stratified random sampling
 - (c) cluster random sampling
 - (d) systematic random sampling
- 1-b. Population includes all the elements from the data set and measurable characteristics known as: **(CO1)** 1
- (a) Sample
 - (b) Sample data
 - (c) Parameter
 - (d) None of these
- 1-c. The _____ sum of squares measures the variability of the observed values around their respective treatment means. **(CO2)** 1
- (a) Treatment

- (b) Error
- (c) interaction
- (d) None of these
- 1-d. For the given the following data pairs (x, y); (1, 1.24), (2, 5.23), (3, 7.24), (4, 7.60), (5, 9.97), (6, 14.31), (7, 13.99), (8, 14.88), (9, 18.04), (10, 20.70), the regression equation is **(CO2)** 1
- (a) $y = 0.490x - 0.053$
- (b) $y = 2.04x$
- (c) $y = 1.98x + 0.436$
- (d) $y = 0.49 \times 17$.
- 1-e. Standard error is the standard deviation of the sampling distribution of an: **(CO3)** 1
- (a) Estimate
- (b) Estimation
- (c) Estimator
- (d) Error of estimation
- 1-f. The connection between a sufficient statistic and an MLE is: **(CO3)** 1
- (a) A sufficient statistic is always an MLE
- (b) There is no connection in general
- (c) All MLE's are linear combinations of sufficient statistics
- (d) If an MLE is unique, then it must be a function of a sufficient statistic
- 1-g. Which of these is a correct alternative hypothesis for a two - tailed test? **(CO4)** 1
- (a) $H_a: \mu \neq 7$
- (b) $H_a: \mu < 7$
- (c) $H_a: \mu > 7$
- (d) None of these
- 1-h. $1 - \alpha$ is the probability associated with: **(CO4)** 1
- (a) Type-I error
- (b) Type-II error
- (c) Level of confidence
- (d) None of the above
- 1-i. A time series is a set of data recorded **(CO5)** 1
- (a) Periodically

- (b) Weekly
- (c) successive points of time
- (d) all the above

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

- (a) $y=T+S+C \times I$
- (b) $y=T+S \times C \times I$
- (c) $y=T+S+C+I$
- (d) $y=T+S \times C+I$

2. Attempt all parts:-

- 2.a. What is the easiest way to sample randomly? (CO1) 2
- 2.b. What is the purpose of using ANOVA? (CO2) 2
- 2.c. How an estimator differs from estimate? (CO3) 2
- 2.d. What is the relation between the width and height of the curve? (CO4) 2
- 2.e. Briefly introduce the Time Series Forecasting. (CO5) 2

SECTION B

30

3. Answer any five of the following:-

- 3-a. A die is thrown 9,000 times and a throw of 3 or 4 is observed 3,240 times. Show that the die cannot be regarded as an unbiased one and find the limits between which the probability of a throw of 3 or 4 lies. **(CO1)** 6
- 3-b. What are the two methods of taking simple random samples? **(CO1)** 6
- 3-c. Calculate the correlation coefficient by Karl Pearson's method between the following data : **(CO2)** 6

x	5	9	13	17	21
y	12	20	25	33	35

3-d. 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 Physic s	1	2	3	4	5	6	7	8	9	10
---------------------------	---	---	---	---	---	---	---	---	---	----

- 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 do the p, d, and q in the ARIMA model mean? **(CO5)** 6

SECTION C

50

4. Answer any one of the following:-

- 4-a. Obtain the expressions for the standard error of sampling distributions of : 10
 (i) sample mean
 (ii) sample variance,
 in random sampling from a large population. Assume that n, the sample size, is large. **(CO1)**
- 4-b. Define sampling and various types of sampling with example. **(CO1)** 10

5. Answer any one of the following:-

- 5-a. The following regression equations and variances are obtained from a correlation table $20x-9y-107=0$, $4x-5y+33=0$, variance of $x = 9$. 10
 Find (i) the mean values of x and y , (ii) the standard deviation of y . **(CO2)**
- 5-b. The following table represents the no. of units of a commodity produced by 3 10
 different workers using 3 different machines:

Workers	A	B	C
Machines			
X	16	64	40
Y	56	72	56
Z	12	56	28

Test:

- (i) whether the mean productivity is the same for the different machines types.
 (ii) whether the three workers differ with regard to mean productivity. **(CO2)**

6. Answer any one of the following:-

- 6-a. A random sample of $n = 6$ has the element 7, 11, 12, 13, 18 and 20. Compute a 10

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. 10
What is the application of sufficiency in estimation? **(CO3)**

7. Answer any one of the following:-

- 7-a. A large hospital hires most of its doctors from the two major universities. Over 10
the last year, 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)**

Uni vers ity A	99	83	89	64	98	85	61	79	91	87	88	
Uni vers ity 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 10
with standard 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 one of the following:-

- 8-a. Explain ARIMA briefly? **(CO5)** 10
- 8-b. Explain the main components of time series data. Which of these would be 10
most prevalent in data relating to unemployment. **(CO5)**