ASSIGNMENT BOOKLET
Bachelor's Degree Programme
(B.Sc./B.A./B.Com.)

STATISTICAL TECHNIQUES

Valid from 1st January, 2023 to 31st December, 2023

- **It is compulsory to submit the Assignment before filling in the Term-End Examination Form.**

- **It is mandatory to register for a course before appearing in the Term-End Examination of the course. Otherwise, your result will not be declared.**

For B.Sc. Students Only

- **You can take electives (56 or 64 credits) from a minimum of TWO and a maximum of FOUR science disciplines, viz. Physics, Chemistry, Life Sciences and Mathematics.**

- **You can opt for elective courses worth a MINIMUM OF 8 CREDITS and a MAXIMUM OF 48 CREDITS from any of these four disciplines.**

- **At least 25% of the total credits that you register for in the elective courses from Life Sciences, Chemistry and Physics disciplines must be from the laboratory courses. For example, if you opt for a total of 24 credits of electives in these 3 disciplines, then at least 6 credits out of those 24 credits should be from lab courses.**

ignou
THE PEOPLE'S
UNIVERSITY

School of Sciences
Indira Gandhi National Open University
Maidan Garhi, New Delhi-110068

(2023)

Dear Student,

Please read the section on assignments in the Programme Guide for Elective Courses that we sent you after your enrolment. A weightage of 30 per cent, as you are aware, has been earmarked for continuous evaluation, **which would consist of one tutor-marked assignment** for this course. The assignment is in this booklet.

Instructions for Formatting Your Assignments

Before attempting the assignment please read the following instructions carefully.

1) On top of the first page of your answer sheet, please write the details exactly in the following format:

ROLL NO.: …………………………………………

NAME: …………………………………………

ADDRESS: …………………………………………

…………………………………………

…………………………………………

COURSE CODE: …………………………….

COURSE TITLE: …………………………….

ASSIGNMENT NO.: ……………………….…

STUDY CENTRE: …………………….…….. **DATE:** ……………………….………………...

PLEASE FOLLOW THE ABOVE FORMAT STRICTLY TO FACILITATE EVALUATION AND TO AVOID DELAY.

2) Use only foolscap size writing paper (but not of very thin variety) for writing your answers.

3) Leave 4 cm margin on the left, top and bottom of your answer sheet.

4) Your answers should be precise.

5) While solving problems, clearly indicate which part of which question is being solved.

6) This assignment is to be submitted to the Study Centre as per the schedule made by the study centre. **Answer sheets received after the due date shall not be accepted.**

   We strongly suggest that you retain a copy of your answer sheets.

7) This assignment is valid only upto December, 2023. If you have failed in this assignment or fail to submit it by December, 2023, then you need to get the assignment for the year 2024 and submit it as per the instructions given in the programme guide.

8) **You cannot fill the Exam Form for this course** till you have submitted this assignment. So solve it and **submit it to your study centre at the earliest.**

We wish you good luck.

Assignment

1. a)   Find a 99% confidence interval for the percentage of copper content of brass, if a sample of size 9 gives the following data: 65, 65, 64, 63, 65, 63, 66, 62, 63.

    Assume a normal population. You may like to use the values given at the end of the question paper. (5)

    b)   Plot a control chart for the mean of data from 12 samples of size 5 each on the diameter of small cylinders in mm. (5)

Sample No.	1	2	3	4	5	6
Mean $(\overline{X})$	4.08	4.112	4.084	4.088	4.108	4.100
Range (R)	0.04	0.02	0.06	0.06	0.04	0.04

7	8	9	10	11	12
4.088	4.096	4.100	4.104	4.140	4.152
0.06	0.04	0.08	0.10	0.004	0.02

2. a)   In the past, the standard deviation of weights of 100 gm packages of tea leaves filled by a machine was 0.8 gm. A sample of 20 packages was drawn and its s.d. was found to be 1.0 gm. Assuming normality at 5% level of significance, test the hypothesis that the population s.d. $\sigma = 0.8$ against the alternative hypothesis that s.d. $\sigma > 0.08$. (4)

    b)   Find the correlation coefficient for data with 20 pairs of values of x and y, if it is given that

$$\sum x_i = 200, \qquad \sum y_i = 500, \qquad \sum x_i^2 = 2900, \qquad \sum y_i^2 = 14100,$$

$$\sum x_i y_i = 6050. \tag{3}$$

    c)   Explain the following terms of quality control with suitable examples: (3)

        i)   Process Capability
        ii) Tolerance

3. a)   i)   Write the sample space of the experiment of drawing 3 screws from a box of right-handed (R) and left-handed (L) screws. (4)

        ii) If event A: At least one R is drawn
             event B: At least one L is drawn
             event C: Exactly 2 right-handed screws are drawn
             event D: Exactly 2 left-handed screws are drawn

             Are the events A and B mutually exclusive? Are C and D mutually exclusive? Justify your answer.

b) The sick leave time for workers in a factory is normal with mean 1000 hours and s.d. 100 hours, per month. How much time t should be budgeted next month exceeds the time corresponding to the probability 20%. (2)

c) Find the moving averages of length 3 for the following data. Also, plot these averages. (4)

Month	Production (Units)
1	140
2	166
3	183
4	253
5	247
6	241
7	242
8	223
9	295
10	450
11	265
12	331

4. a) Find the mean and variance of the following data: (6)

Class	0 – 20	20 – 40	40 – 60	60 – 80	80 – 100	100 – 120	120 – 140
Frequency	1	3	10	15	12	6	3

b) Suppose from a total of 120 guava bearing trees in a village, 5 clusters of 4 trees each are selected and yield (in kg) recorded is a given in the following table: (4)

Cluster Number	Yield (in Kg)			
	Tree 1	Tree 2	Tree 3	Tree 4
1	2	22	8	6
2	26	10	19	11
3	4	4	2	15
4	10	2	4	7
5	9	15	10	10

Estimate the average yield (in kg) per tree of guava using cluster sampling method and simple random sampling method.

5. a) A sample of 6 boxes is to be selected from 30 boxes of mangoes numbered 1 to 30. Find the samples using: (6)

i) Direct WR method
ii) Remainder approach
iii) Quotient approach, stating clearly the starting random numbers chosen by you.

b) Find the regression line of y on x (4)

x	4	6	8	10
y	2.3	4.1	5.7	6.9

6. a)   A random sample of size 4 is taken from each of 3 independent normal random
         variables X_1, X_2, X_3 resulting in the following data: (6)

X_1	13	11	16	22
X_2	16	8	21	11
X_3	15	12	25	10

    Assuming that the 3 variables have equal variances, test at the 0.05 significance level,
    the hypothesis that X_1, X_2, X_3 have the same mean using ANOVA. You may like to
    use following values.

   b)   Suppose you are waiting at a bus stand for a bus. Every 10 minutes, you are recording
        the number of buses that pass by you. Data recorded is as follows: (4)

Time	0 – 10 min	10 – 20 min	20 – 30 min	30 – 40 min	40 – 50 min	50 – 60 min
Number of buses	1	2	0	5	4	2

    Identify the distribution, the buses follow. Calculate the average number of buses that
    pass you and the variance.

7. a)   In the amount of cosmic radiations to which a person is exposed while flying across a
         specific continent is a normal random variable with mean 4.35 units and S.D. 0.59
         units. Find the probabilities that the amount of exposure during such a flight is (i)
         between 4.00 and 5.00 units, (ii) at least 5.50 units.   (4)

   b)   A sampling random sample of size 100 has mean 15 and population variance 25. Find
        an interval estimate of the population mean with a confidence level of 99% and 95%. (4)

   c)   A hypothetical population consists of the numbers 2, 5 and 7. Write all possible simple
        random samples of size 2 (with replacement). Verify that the sample mean is an
        unbiased estimator of the population mean.   (2)

8. a) The following random samples are measurements of the heat producing capacity in
       millions of calories per ton of specimens of coal from mines:

Mine I	Mine II
8260	7950
8130	7890
8350	7900
8070	8140
8340	7920
	9.5

    Test at 5% level of significance whether the difference between the means of two samples
    is significant?   (4)

   b) A company operates four machines on three separate shifts daily. The following table
      presents the data for machine break-downs resulted during a 6 month time period:

5

Shift	Machine				Total
	A	**B**	**C**	**D**	
1	10	12	6	7	35
2	10	24	9	10	53
3	13	20	7	10	50
Total	33	56	22	27	138

At 5% level of significance test the hypothesis that for an arbitrary breakdown in the machine causing the breakdown and the shifts are independent. (6)

9. a) The three drying techniques for curing a glue were studied and the following times were observed:

Formula A	Formula B	Formula C
13	13	4
10	11	1
8	14	3
11	14	4
8		2
		8

At $\alpha = 0.05$, test the hypothesis that the average times for three formulae are same. (6)

b) The following data represent the number of defects discovered at a factory on 20 successive batches of 10 cars each:

Batch No.	No. of Defects
1	143
2	160
3	150
4	105
5	98
6	74
7	85
8	95
9	76
10	68
11	63
12	87
13	90
14	81
15	94
16	68
17	95
18	90
19	93
20	73

Does it appear that the production process is in statistical control throughout? Give justification. (4)

10. Which of the following statements are true and which are false? (10)

   a) If X is continuous random variables following uniform distribution with values in $[2,6]$, then $P(X = 4) = 0$

   b) A 90% confidence interval calculated using a sample of size 40 is smaller than the one calculated using a sample of size 100.

   c) If a dice is rolled twice, then the number of possible outcomes is 36.

   d) If every 5th student on the roll is selected and her height is measured, the sampling method used is simple random sampling.

   e) A type I error is committed when we fail to reject a null hypothesis when the alternative hypothesis is true.