

ASSIGNMENT BOOKLET

Post Graduate Diploma in Applied Statistics (Specialisation in Industrial Statistics)

MST-001 to MSTL-002

(Valid from 1st January, 2021 to 31st December, 2021)

**It is compulsory to submit the assignments
before filling the Examination Form.**



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

Dear Student,

Please read the information on assignments in the Programme Guide that we have sent you after your enrolment. A weightage of 30%, as you are aware, has been earmarked for continuous evaluation, **which would consist of one tutor-marked assignment** for this course. The assignments for MST-001 to MSTL-002 have been given 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:

ENROLLMENT NO :.....
NAME :.....
ADDRESS :.....
.....
.....

PROGRAMME CODE:

COURSE CODE:

COURSE TITLE:

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) This assignment is to be submitted at the Study Centre.

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

- 6) This assignment is valid up to December 31, 2021.
- 7) **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**. If you wish to appear in the **TEE, June 2021**, you should submit your TMAs by **March 31, 2021**. Similarly, If you wish to appear in the **TEE, December 2021**, you should submit your TMAs by **September 30, 2021**.

We wish you good luck.

TUTOR MARKED ASSIGNMENT

MST-001: Foundation in Mathematics and Statistics

Course Code: MST-001

Assignment Code: MST-001/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are **True** or **False**. Give reason in support of your answer: (5×2=10)

(a) Intersection of set of rational and irrational numbers is non-empty.

(b) The function $|x - 2|$ is differentiable everywhere on the real line.

(c) The trace of the matrix $\begin{bmatrix} 3 & 5 & 8 \\ -1 & 7 & 9 \\ 0 & 0 & 5 \end{bmatrix}$ is 130.

(d) Six persons can sit on six chairs placed around a circular table in 120 numbers of ways.

(e) The range of the data shown in the following frequency distribution is 70.

Classes	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequencies	2	7	3	8	4	1	0

2 (a) An institution made a policy for delay of payment beyond a certain date to be paid by its customers. Penalty as per policy are as follows: Rs 50 for the first day, Rs 100 for the second day, Rs 150 for the third day, and so on. If a customer makes the payment on a delay of 31 days then how much he/she has to pay as penalty charges. (4)

(b) If $A = \{3, 6, 9, 12, \dots\}$ and $B = \{15, 30, 45, \dots\}$ then verify whether two sets A and B are equivalent or not. (3)

(c) Find the sum $-7 + (-12) + (-17) + \dots + (-507)$. (3)

3. (a) Express $5.10.15.20.25.30.35$ in terms of factorial.

(b) How many different signals are possible with 4 blue, 3 red, 2 white and 2 green flags by using all at a time in a queue?

(c) If in a hall there are 10 randomly selected students then how many numbers of ways are there such that all of them have different birthday. Assume that all of them have their birth day in non-leap years. (2+4+4)

4. Discuss the continuity and differentiability of the following function at $x = 5$. (5+5)

$$f(x) = \begin{cases} |x - 5|, & x \neq 5 \\ 0, & x = 5 \end{cases}$$

5. Evaluate the following integrals:

i) $\int x^5 e^{x^3} dx$ ii) $\int_{-4}^3 |x^2 + x - 6| dx$ (5+5)

6. The cost of 2 pens, 3 note-books, and 4 books is Rs 260. The cost of 4 pens, 5 note-books, and 3 books is Rs 290. The cost of 3 pens, 2 note-books, and 5 books is Rs 320. Find the cost of 3 pens, 2 note-books and 4 books. You are bound to use the matrix techniques to solve the given equations. **(10)**

7. (a) If $A = \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & 8 \\ 7 & 9 \end{bmatrix}$ then verify that $(AB)^{-1} = B^{-1}A^{-1}$. **(10)**

- (b) Distinguish characteristic features of each of the four measurement scales: (I) Nominal (II) Ordinal (III) Interval, and (IV) Ratio Scale. Name the measurement scale for each of the following variable.

- i) Waiting time for a bus at a particular bus stop in a certain city.
- ii) Number of pages in a book.
- iii) Temperature in degree Celsius.
- iv) Classification of cancer patients according to their severity of diseases.
- v) Monthly income of the family
- vi) Saving (Income in that particular month – expenditure in the same month) of the family. Keep in mind that expenditure may be more than income in that particular month of some family(ies)
- vii) Temperature of water by touch: cold, lukewarm and hot
- viii) Height of students of a class
- ix) State of residence of each student of PGDAST programme
- x) Balance of Smart card of Delhi Metro hold of different passengers

(10×1=10)

8. (a) Compare the following two ratio scale data sets using suitable graphical analysis tool:

Score before training: 12, 17, 15, 23, 27, 21, 24, 19, 20, 23, 16

Score after training: 16, 22, 20, 24, 30, 20, 30, 17, 27, 31, 19

Also interpret the finding on the basis of the graph.

- (b) Comment on the height of a bar in a bar diagram and height of a bin in a histogram. Are two heights have the same meaning? Explain it with the help of an example.

(10+10)

TUTOR MARKED ASSIGNMENT

MST-002: Descriptive Statistics

Course Code: MST-002

Assignment Code: MST-002/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are true or false and also give the reason in support of your answer: (5×2=10)

- a) If standard deviation of x is 5, standard deviation of $y = 2x - 3$ is 7.
- b) Mean deviation is least when calculated from the median.
- c) The correlation coefficient between x and $(a - x)$ is -1 .
- d) The regression coefficients b_{yx} and b_{xy} of a data are 1.2 and 0.8, respectively.
- e) If $(AB) = 10$, $(\alpha B) = 15$, $(A\beta) = 20$ and $(\alpha\beta) = 30$ then A and B are associated.

2. a) Compute coefficient of variation for factory A and B from the following:

Daily Wages (Rs.) :	12	15	17	22	25	30
Workers in Factory A:	5	10	20	35	12	8
Workers in Factory B:	15	25	30	35	21	19

Find: (i) Which factory pays higher average daily wages? (ii) In which factory are wages more variable?

- b) The mean and the standard deviation of a characteristic of 100 items were found to be 60 and 10 respectively. At the time of calculations, two items were wrongly taken as 5 and 45 instead of 30 and 20. Calculate the corrected mean and corrected standard deviation.

(10+6)

3. Calculate the first four moments of the following distribution about the mean and hence find β_1 and β_2 .

X:	0	1	2	3	4	5	6	7	8
F:	1	8	28	56	70	56	28	8	1

(14)

4(a) Fit an exponential curve of the form $Y = ab^X$ to the following data:

X:	1	2	3	4	5	6	7	8
Y:	1.0	1.2	1.8	2.5	3.6	4.7	6.6	9.1

(b) In a partially destroyed laboratory, record of an analysis of correlation data, the following results only are legible:

Variance of $X = 9$. Regression equations: $8X - 10Y + 66 = 0$, $40X - 18Y = 214$.

What are: (a) the mean values X and Y , (ii) the correlation coefficient between X and Y , and (iii) the standard deviation of Y ? (8+12)

- 5 a) The following table gives, according to age, the frequency of marks obtained by 100 students in an intelligence test:

Age in years → Marks ↓	18	19	20	21	Total
10-20	4	2	2	–	8
20-30	5	4	6	4	19
30-40	6	8	10	11	35
40-50	4	4	6	8	22
50-60	–	2	4	4	10
60-70	–	2	3	1	6
Total	19	22	31	28	100

Calculate the Coefficient of Correlation between the Age and Marks. (14)

- b) An individual purchases three qualities of pencils. The relevant data are given below.

Quality	Price per pencil (Rs.)	Money spent (Rs.)
A	1.00	50
B	1.50	30
C	2.00	20

Calculate the average price per pencil. (06)

- 6(a) An investigation of 23,713 households was made in an urban and rural mixed locality. Of these 1,618 were farmers, 2,015 well-to-do and 770 families were having at least one graduate. Of these graduate families 335 were those of farmers and 428 were well-to-do, also 587 well-to-do families were those of farmers and out of them only 156 were having at least one of their family members as graduate. Obtain all the ultimate class frequencies.

- (b) Investigate the association between darkness of eye-colour in father and son from the following data:

Fathers with dark eyes and sons with dark eyes: 50

Fathers with dark eyes and sons with not dark eyes: 79

Fathers with not dark eyes and sons with dark eyes: 89

Fathers with not dark eyes and sons with not dark eyes: 782

Also tabulate for comparison the frequencies that would have been observed had there been no heredity. (8+12)

TUTOR MARKED ASSIGNMENT

MST-003: Probability Theory

Course Code: MST-003

Assignment Code: MST-003/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are **True** or **False** and also give the reason in support of your answer. (5×2=10)

(a) If $P(A^c) = \frac{1}{4}$, $P(B^c) = \frac{3}{16}$ and $P(A \cup B) = \frac{7}{8}$, then $P(A \cap B) = \frac{5}{16}$.

(b) Two cards are drawn one at a time from a 52-card standard deck. Then the probability that the second card is red, if the drawing is done with replacement is $\frac{1}{2}$.

(c) If standard deviation of a random variable X is 3 and $Y = -5X - 10$ then $SD(Y) = -85$.

(d) If X is a continuous random variable then $P[a < X < b] \neq P[a \leq X < b]$.

(e) If a random variable X may assume values $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$ then it cannot be discrete random variable.

2. Roll two fair dice, one red and one blue, and consider the events:

A: "The red die lands on 4",

B: "The sum on the dice is 9" and

C: "The blue die lands on an odd number".

Determine which pairs of events are independent. (10)

3. (a) An urn contains 10 white marbles, 15 blue marbles, and 20 red marbles. Five marbles are selected, one at a time, with replacement. Find the probability that at least one color will be missing from the 5 selected marbles.

(b) There are three coins in a box. When tossed, one of the coins comes up heads only 30% of the time, one of the coins is fair, and the third comes up heads 80% of the time. A coin is selected at random from the box and tossed three times. If two heads and a tail come up in this order (HHT) what is the probability that the coin was the fair coin? (5+5)

4. (a) A fair die is tossed. If the resulting number is even, you add 1 to your score and get that many rupees. If the resulting number is odd, you add 2 to your score and get that many rupees. If X denotes the random variable counting your gain in rupees, then write probability distribution of X . (5)

(b) A hand of three cards is chosen from a standard deck of 52 cards. You get Rs 3 for each heart in your hand. If X denotes the random variable measuring your gain, then find expected value of X . (5)

5. A coin is biased so that heads appears with probability $2/3$ and tails with probability $1/3$. This coin is tossed three times. If X denotes the number of heads occurring and Y denotes the number of tails that occur up to the first head, if any, that appears, then determine
- the joint probability mass function of X and Y .
 - marginal probability mass functions of the random variables
 - $P[X \leq 1 | Y \leq 1]$ **(6+2+2)**
6. (a) In a certain book, there is one misprint per two pages, on the average. What is the probability that there are two or more misprints on a randomly chosen page? **(5)**
- (b) An insurance company models the number of days elapsing between the beginning of a calendar year and the moment a high-risk driver gets into an accident by means of an exponential random variable with parameter λ . If they expect that 10% of the high-risk drivers will get into an accident during the first 30 days of the calendar year, find the probability that a high-risk driver will get into an accident during the first 40 days of a calendar year. **(5)**
7. (a) A drunk has five keys on his key chain, and only one will open the front door of his house. He tries each key until he finds the right one. Assume that he is so drunk that he may repeat the wrong key any number of times. On average how many trials he will make to open the front lock of his house. **(10)**
- (b) In a certain game, Ravi beats Pawan with probability $2/5$. They play a series of games until one of them wins six games. Assume that each game is independent of the other. Find the probability that the series will end after 10 games. **(10)**
8. (a) A hospital specialises in heart surgery. During the year 2018-19, 2000 patients were admitted for treatment. Average payment made by a patient was Rs. 120000 with a standard deviation of Rs 25000. Assuming the distribution of payments to normal, find:
- The number of patients who paid between Rs. 1,00,000 and Rs. 1,75,000.
 - The probability that a patient's bill exceeds Rs. 75,000
 - The maximum amount paid by the lowest paying one-third patient.
- (b) The menstrual cycle in woman following normal distribution has a mean of 28 days and S.D. of 2 days. How frequently would you expect a menstrual cycle of
- More than 30 days
 - Less than 22 days

(10+10)

TUTOR MARKED ASSIGNMENT

MST-004: Statistical Inference

Course Code: MST-004

Assignment Code: MST-004/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are **True** or **False**. Give reason in support of your answer: (5×2=10)

- (a) If the probability density function of a random variable X follows F-distribution is

$$f(x) = \frac{1}{(1+x)^2}, x \geq 0$$

then degrees of freedom of the distribution will be (2,2).

- (b) If X_1, X_2, X_3, X_4 and X_5 is a random sample of size 5 taken from an Exponential distribution, then estimator T_1 is more efficient than T_2 .

$$T_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}, T_2 = \frac{X_1 + 2X_2 + 3X_3 + 4X_4 + 5X_5}{15}$$

- (c) If a researcher wants to half the standard error of the estimator of the mean then he/she will increase the sample size 4 times.

- (d) For testing the pulse rate of one month old child is to be 120 per minute, a researcher measures (per minute) the pulse rate of 10 one month old children and applies sign test and Wilcoxon signed-rank test, then power of the Wilcoxon signed-rank test is more than the sign test.

- (e) The Z calculated value for testing the following hypothesis is 1.56.

$$H_0 : \mu = 100 \text{ and } H_1 : \mu \neq 100$$

Then the p-value will be 0.256.

- 2 The lifespan (in '000 hours) of four LED bulbs of 07 watts are as follows:

40,46,48, 50

- (i) Write all samples of size 2 without replacement from the above observations. Compute the sampling distribution and find mean and standard error of the distribution.

- (ii) Write all samples of size 3 without replacement from the above observations. Compute the sampling distribution and find mean and standard error of the distribution.

- (iii) Compare means and standard errors and draw the conclusion. (4+4+2)

3. An ambulance agency claims that the mean length of service times is 10 minutes with standard deviation 3 minutes. An investigator suspects that this claim is wrong. She takes a random sample of 20 services and finds the mean length of service times is 12 minutes with a SD 5 minutes. Assume that the service time of the ambulance follows normal distribution.

- (i) What is the probability that the mean length of service times is less than and equal to 12 minutes?

- (ii) Find 99% confidence interval for standard deviation of the mean length of service times.

(5+5)

4. The department of transportation has mandated that the average speed of cars on interstate highways be no more than 70 km per hour in order. To check that the people follow it or not, a researcher took a random sample of 186 cars and found that the average speed was 68 km per hour with a standard deviation 0.6 km per hour.
- Find the standard error of the mean
 - Construct the interval around the sample mean that would contain the population mean 95% of the time.
 - If the researcher wants to test that the true mean speed on its highways is 70 km per hour or less with 95% confidence then
 - State null and alternative hypotheses
 - Name the test which is suitable in this situation and why?
 - Calculate the value of test statistic and critical value.
 - Draw the conclusion on the basis of the applied test. (1+3+6)
5. (a) In a trial, the null hypothesis is that an individual is innocent of a certain crime. Would the legal system prefer to commit a type-I or type-II error with this hypothesis?
- (b) A coal-fired power plant is considering two different systems for pollution abatement. The first system has reduced the emission of pollutants to acceptable levels 68 percent of the time, as determined from 200 air samples. The second, more expensive system has reduced the emission of pollutants to acceptable levels 76 percent of the time, as determined from 250 air samples. If the expensive system is significantly more effective than the inexpensive system in reducing pollutants to acceptable levels, then the management of the power plant will install the expensive system. To decide that
- State null and alternative hypotheses
 - Name the test which is suitable in this situation and why?
 - Calculate the value of test statistic and critical value.
 - Which system will be installed if management uses a significance level 0.01 in making decision? (3+7)
6. A newspaper publisher, trying to pinpoint his market's characteristics wondered whether newspaper readership in the community is related to readers' educational achievement. A survey questioned adults in the area on their level of education and their frequency of readership. The results are shown in the following table:

Frequency of Readership	Level of Educational Achievement			
	Professional or postgraduates	College graduates	High school grad	Did not complete high school
Never	10	17	11	21
Sometimes	12	23	8	5
Morning or evening	35	38	16	7
Both editions	28	19	6	13

Do these data provide the sufficient evidence that frequency of newspaper readership in the community differ according to the reader's level of education?

- (i) Calculate expected frequencies for this problem
- (ii) State null and alternative hypotheses.
- (iii) Calculate the value of the test statistic.
- (iv) At the 5% significance level, what is your conclusion? (3+2+3+2)

7. If magnitude of earthquakes recorded in a region of a country follows a distribution with parameter θ whose pdf is given below:

$$f(t) = \frac{1}{\theta^2} t e^{-t/\theta}, t > 0, \theta > 0$$

then

- (i) Show that the estimators of the parameter θ using maximum likelihood and method of moments are same,
 - (ii) Show that maximum likelihood estimator is unbiased and sufficient for the parameter θ . (10+10)
8. A chemist developing insect repellents wishes to know if a newly developed formula gives greater protection from insect bites than that given by the leading product on the market. In the experiment, 14 volunteers each had one arm sprayed with the old product and the other sprayed with the new formula. Then each subject placed his arms into two chambers filled with equal number of mosquitoes, gnats and other biting insects. The numbers of bites received on each arm are as follows:

Subject	Old formula	New formula	Subject	Old formula	New formula
1	5	3	8	4	2
2	2	1	9	2	5
3	5	5	10	6	2
4	4	1	11	5	3
5	3	1	12	7	3
6	6	4	13	4	1
7	2	4	14	3	2

To test the new formula is more effective than the old one:

- (i) State null and alternative hypotheses
- (ii) Can you apply both parametric and non-parametric tests in this problem and why?
- (iii) Write the assumptions to apply the suitable parametric test.
- (iv) Apply the parametric test by assuming the assumptions write in part (iii) are fulfilled and write the conclusion.
- (v) Apply the non-parametric test and write the conclusion.
- (vi) Compare the conclusions drawn in parts (iv) and (v). (2+2+2+6+6+2)

TUTOR MARKED ASSIGNMENT

MST-005: Statistical Techniques

Course Code: MST-005

Assignment Code: MST-005/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are true or false and also give the reason in support of your answer: (2×5=10)

- a) $V_{opt}(\bar{x}_{st})$ lies between $V_{prop}(\bar{x}_{st})$ and $V_{Random}(\bar{x}_{st})$.
- b) In cluster sampling the variance within clusters is greater than the variance between clusters.
- c) While analysing the data of a 5×5 Latin Square design the error d.f. is equal to 12.
- d) In a Two way analysis of variance with 4 blocks & 4 treatments the degree of freedom for the total variation is 15.
- e) Suppose a random number generated by Middle. Square Method is 16, then the next random number will be 44.

- 2(a) Explain by considering a sample of size $n = 2$ from a population consisting of five members 2, 3, 6, 8, 11 that SRSWOR gives a better estimator of population mean than SRSWR. (6)

- (b) From the following data find mean and variance of the stratified sampling under (i) proportional and (ii) Neyman Allocation

Strata	N_i	n_i	S_i	\bar{y}_i
1	80	29	12	80
2	160	39	8	30
3	260	32	4	10
	500	100		

(6)

- 3(a) In a population of size $N = 5$, the values of the population characteristics are 1, 3, 5, 7, 9, a sample of size 2 is drawn. Verify that \bar{y} is an unbiased estimate of \bar{Y} and $V(\bar{y})$ is equal to

$$V(\bar{y}) = \frac{N-n}{N.n} S^2. \quad (10)$$

- (b) In order to compare the mileage yields of 3 kinds of gasoline, several tests were run and the following results were obtained:

Gasoline A:	19	21	20	18	21	21
Gasoline B:	23	20	22	20	24	23
Gasoline C:	20	17	21	19	20	17

Carry out the analysis of variance test and test whether the observed differences between the means of 3 kinds of gasoline may be attributed to chance at 5% level of significance.

(7)

4. A manurial trial with six levels of Farm Yard. Manure (FYM) was carried out in a randomised block design with 4 replications at the experimental station Junagarh with a new study the rate of decomposition of organic matters in soil and its synthetic capacity in soil on cotton crop. The yield per plot in kg for different levels of FYM and replications is given below:

Cotton Yield Per Plot (in Kg)

Levels of FYM	Replications			
	I	II	III	IV
1	6.90	4.60	4.40	4.81
2	6.48	5.57	4.28	4.45
3	6.52	7.60	5.30	5.30
4	6.90	6.65	6.75	7.75
5	6.00	6.18	6.50	5.50
6	7.90	7.57	6.80	6.62

Carry out the Analysis of variance test and draw the conclusions. (10)

5. In the following data two values are missing. Estimate these values and analyse the design given as:

Treatment	Blocks		
	I	II	III
A	12	14	12
B	10	y	8
C	x	15	10

(7)

6. The following data is the data pertaining to a feeding trial on sheep. Treatments
A: Grazing only
B: Grazing + Maize Supplements
C: Grazing + Maize + Protein Supplement P₁
D: Grazing + Maize + Protein Supplement P₂
E: Grazing + Maize + Protein Supplement P₃

Layout and Wool Yield (100 gm) is given as:

32(D)	33(E)	30(C)	28(B)	24(A)
51(C)	45(D)	41(A)	45(E)	29 (B)
41 (E)	29 (A)	24 (B)	36 (D)	35 (C)
38 (B)	39(C)	42(E)	23(A)	37(D)
38(A)	24(B)	21(D)	29(C)	26(E)

- i) Analyse the design with appropriate method and calculate the CD for the treatment mean yield.
ii) Calculate the relative efficiencies of the above L.S.D over R.B.D and C.R.D.

7. For a factorial experiment with 3 factors, N, P and M each at two levels, the design and yield per plot are given below. Analyse the experiment.

Block – 1	np	nm	pm	(1)
	48.81	58.88	46.11	38.62
	m	p	npm	n
	40.49	32.75	61.55	55.07

Block – 2	np	pm	(1)	nm
	50.43	52.31	40.26	49.62
	m	p	npm	n
	32.36	51.94	48.89	53.86

Block – 3	n	npm	p	m
	47.37	46.87	37.25	46.94
	pm	(1)	nm	np
	39.30	39.23	49.93	51.43

(14)

- 8(a) Times between successive crashes of a computer system were generated for a 6-months period and are given in increasing order as follows (time in hours):
1, 10, 20, 30, 40, 52, 63, 70, 80, 90, 100, 102, 130, 140, 190, 210, 266, 310, 530, 590, 640, 1340.

The value of parameter α and Mean ($1/\alpha$) is given as 0.00435 and 230 hrs, respectively.

Use kolmogrov-Smirnov test to examine the goodness of fit of exponential distribution.

- (b) Generate ten uniform random numbers $U(0, 1)$ from the multiplicative LCG given below:

$$x_i = (5x_{i-1}) \bmod 32 \text{ with } x_0 = 1$$

Also obtain a sequence of heads and tails using above generated random numbers.

(10+10)

TUTOR MARKED ASSIGNMENT

MSTE-001: Industrial Statistics-I

Course Code: MSTE-001

Assignment Code: MSTE-001/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are **True** or **False**. Give reason in support of your answer: (5×2=10)
 - (a) Twenty pieces of different length of cloth contained 2, 4, 1, 3, 5, 4, 2, 7, 3, 5, 2, 2, 4, 5, 6, 4, 2, 1, 2, 4 defects respectively. To check the process is under control with respect to the number of defects, we should use p-chart.
 - (b) If density function of the time to failure of an appliance is $f(t) = \frac{32}{(t+4)^3}$; $t > 0$ then reliability of the appliance for two years will be 0.444.
 - (c) If the probabilities are associated with the occurrence of different states of nature, then the situation is known as decision making under uncertainty.
 - (d) If there are 1% defectives notes of Rs.500 in a lot of 10000 notes of Rs 500, then the lot quality will be 99%.
 - (e) A system has four components connected in parallel configuration with reliability 0.2, 0.5, 0.4, 0.6. To improve the reliability of the system, we have to replace the weakest component to the more reliable component.
2. (a) A company manufactures water pumps. The quality control inspector of the company takes a sample of 100 water pumps at regular intervals. The numbers of defective pumps for 15 samples are given below:

Sample No.	Defective Pumps	Sample No.	Defective Pumps	Sample No.	Defective Pumps
1	5	6	0	11	6
2	6	7	4	12	1
3	3	8	8	13	10
4	2	9	2	14	2
5	1	10	2	15	1

Use the data to construct a suitable chart. Observe the results and comment on the control of the process as indicated by the chart. (5)

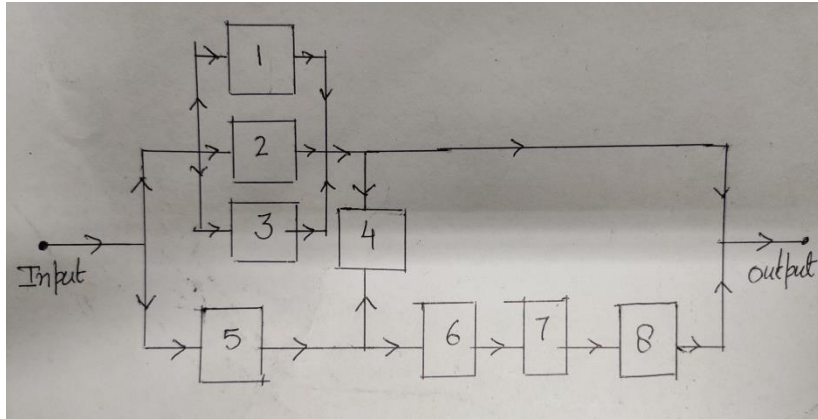
- (b) A restaurant produces fresh burgers for its customers every day. The company is known for supplying fresh burgers and never uses burgers prepared on the previous day. Demand for burgers is uncertain, preparation capacity is limited, and the restaurant has the option of producing 0, 100, 200, 300 and 400 burgers every day. It has been estimated that the cost of producing each burgers pack is Rs.15. Each burger is sold for Rs. 20. Prepare a payoff matrix when 0, 100, 200, 300 or 400 demands of the burgers turn up on any given day. Prepare an opportunity loss table and hence find the optimum strategy. (5)

3. A leather bag manufacturing company supplies bags in lots of size 200 to a buyer. A sample of 2 bags is drawn and the corresponding lot is accepted if and only if both bags are non-defective. The company and the buyer decide that $AQL = 0.04$ and $LTDD = 0.10$. If there are 15 defective bags in each lot, compute the
- Probability of accepting the lot,
 - Producer's risk and consumer's risk,
 - AOQ, if the rejected lots are screened and all defective bags are replaced by non-defectives, and
 - Average total inspection. (2+4+2+2)
4. An office supply company ordered a lot of 400 printers. When the lot arrives the company inspector will randomly inspect 12 printers. If more than three printers in the sample are non-conforming, the lot will be rejected. If fewer than two printers are non-conforming, the lot will be accepted. Otherwise, a second sample of size 10 will be taken. Suppose the inspector finds two non-conforming printers in the first sample and two in the second sample. Also AQL and $LTPD$ are 0.05 and 0.10 respectively. Let incoming quality be 4%.
- What type of acceptance plan is used here?
 - What is the probability of accepting the lot at the first sample?
 - What is the probability of accepting the lot at the second sample? (2+3+5)
5. A person has two independent investments A and B, but he can undertake only one at a time due to certain constraints. He can choose A first and then stop, or if A is successful, then take B or vice-versa. The probability of success of A is 0.6 while for B it is 0.4. Both investments require an initial capital outlay of Rs. 10,000 and both return nothing if the venture is unsuccessful. Successful completion of A will return Rs. 20,000 (over cost) and successful completion of B will return Rs. 24,000 (over cost). Draw an appropriate decision tree and determine the best strategy. (10)
6. Two breakfast food manufacturers X and Y are competing for an increase market share. The payoff matrix, shown in the following table, describes the increase in the market share for X and decrease in the market share of Y.

X	Y			
	Give Coupons	Decrease Price	Maintain Present Strategy	Increase Advertising
Give Coupons	10	-10	20	5
Decrease Price	30	5	60	15
Maintain Present Strategy	-15	10	0	30
Increase Advertising	10	-15	35	5

- Check whether saddle point exist or not.
- If saddle point does not exist then determine optimal strategies for both the manufacturers and value of the game. (2+8)

7. (a) A system has eight independent components and reliability block diagram of it shown blow.



Components 1, 2 and 3 are not identical and at least two components of this group must be available for system success. The reliability of component 1 to 8 (for a mission 1 year) is given below:

$$R_6 = R_7 = R_8 = 0.80, R_1 = 0.60, R_2 = 0.40, R_3 = R_5 = 0.50 \text{ and } R_4 = 0.60.$$

Find reliability of the system.

(10)

- (b) A system having Weibull failure distribution with pdf as

$$f(t) = \begin{cases} \frac{1}{\theta} e^{-t/\theta}; t > 0, \theta > 0 \\ 0 & ; \text{otherwise} \end{cases}$$

then

- i) Compute the reliability function of the system,
- ii) Find the hazard rate,
- iii) Find the MTTF, and
- iv) What is the life of the system if reliability of 0.90 is desired?

(3+2+2+3)

8. Twelve samples of 4 LED bulbs were selected at regular intervals from a LED bulbs manufacturing company. If bulbs have mean life equal to 2000 hours, it is considered satisfactory. The SD of life of the bulbs is expected to be 520 hours. On testing the samples, the failure times (in hours) were recorded and given below:

Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Sample 10	Sample 11	Sample 12
2081	1528	1984	1728	1804	2002	1994	1616	1982	2132	2134	1749
1363	1330	2384	1972	1845	1804	2023	1832	2342	1998	2140	1948
2092	2053	2194	1647	2132	1760	2136	1497	2132	1554	1756	2050
2385	1945	1456	1792	2024	2035	1842	1692	1994	1777	1994	1857

- i) Prepare control chart for mean when the mean life and SD of the life of the LED bulbs are known and draw the conclusion.
- ii) If mean and SD of the life of the LED bulbs are to be unknown, then prepare the control charts for mean and variability. If process is out of control then calculate the revised control limits.

iii) If specification limits as the $2000 \pm SD$, then find the process capability. Does it appear that the manufacturing process is capable of meeting the specification requirements?

(4+12+4)

TUTOR MARKED ASSIGNMENT

MSTE-002: Industrial Statistics-II

Course Code: MSTE-002

Assignment Code: MSTE-002/TMA/2021

Maximum Marks: 100

Note: All questions are compulsory. Answer in your own words.

1. State whether the following statements are true or false and also give the reason in support of your answer. (2×5=10)
 - a) If the arrival rate is 12 per hour and service rate is 4 per hour, then the probability of no customer in queue is 0.3.
 - b) If the coefficient of determination is 0.833, the number of observations and independent variables are 12 and 3, respectively, then Adjusted R^2 will be 0.84.
 - c) The Set $S = \{(x, y) : 0 \leq y \leq 5 \text{ when } 0 \leq x \leq 2 \text{ and } 3 \leq y \leq 5 \text{ when } 2 \leq x \leq 7\}$ is not a convex set.
 - d) The solution to a transportation problem with 3-rows (supplies) and 3-columns (destinations) is feasible if number of positive allocations is 6.
 - e) Variations which occur due to natural forces and operate in a regular and periodic manner over a span of less than or equal to one year are termed as cyclic variations.

2. Solve the following LPP using simplex method:

$$\text{Maximize } Z = 10x_1 + x_2 + 2x_3$$

Subject to the constraints:

$$4x_1 + x_2 - 6x_3 = 7$$

$$6x_1 + x_2 - 4x_3 \leq 5$$

$$3x_1 - x_2 - x_3 \leq 0$$

$$x_1, x_2, x_3 \geq 0 \quad (10)$$

3. Let X is the advertisement expenditures (in Lakh Rs.) and Y is the sales (in Lakh Rs.). Let the data are

X:	1182	1172	1264	1493	1571	1711	1804	1840	1956	1954
Y:	129	135	147	160	171	184	198	223	240	293

Estimate the parameters and find the estimated linear equation. Whether the advertisement influences the sale of product? Test and comment on the goodness of fit of the model. (10)

4. a) The production department for a company requires 3600 kg of raw material for manufacturing a particular item per year. It has been estimated that the cost of placing an order is Rs 36 and the cost of carrying inventory is 25 percent of the investment in the inventories. The price is Rs10 per kg. The purchase manager wishes to determine an ordering policy for raw material.

b) Arrivals at telephone booth are considered to be Poisson with an average time of 10 minutes between arrival and the next. The length of phone call is assumed to be distributed exponentially with mean 3 minutes.

- i) What is the probability that a person arriving at the booth will have to wait?
- ii) The telephone department will install a second booth when convinced that an arrival would expect waiting for at least 3 minutes for phone call. By how much should the flow of arrivals increase in order to justify a second booth?
- iii) What is the average length of the queue that forms from time to time?
- iv) What is the probability that it will take him or her more than 10 minutes altogether to wait for the phone and complete his or her call?

(6+6)

5. The annual sales revenue (in lakhs of Rs) of a product as a function of sales force (number of salesmen) and annual advertising expenditure (in Thousands of Rs) for the past 10 years are summarised in following table:

Annual Sales Revenue Y (in Lakhs)	Sales Force X_1 (in Number)	Annual Advertising Expenditure X_2 (in Thousand)
100	40	140
115	65	115
125	40	190
135	90	80
105	115	100
145	80	140
110	50	115
120	60	150
135	70	130
175	100	160

Obtain a regression model to forecast the annual sales revenue of the product using Matrix Method. (10)

6. A solicitors' firm employs typists on hourly price-rate basis for their daily work. There are five typists and their charges and speed are different. According to an earlier understanding only one job is given to one typist and the typist is paid for a full hour even if he works for a fraction of an hour. Find the least cost allocation for the following data:

Typist	Jobs				
	P	Q	R	S	T
A	85	75	65	125	75
B	90	78	66	132	78
C	75	66	57	114	69
D	80	72	60	120	72
E	76	64	56	112	68

(10)

7. a) Obtain seasonal Indices by the “Moving average” method from the following data:

Quarterly output of a Factory				
Year	I	II	III	IV
2010	65	58	56	61
2011	68	63	63	67
2012	70	59	56	52
2013	60	55	51	58

(12)

(b) For the following Auto regressive model

$$X_t = 0.7X_{t-1} - 0.4X_{t-2} + a_t$$

i) Verify whether the series is Stationary

ii) Obtain $\rho_k : k = 1, 2, 3, 4$ and 5

Plot the Correlogram.

(08)

8. a) A company has three production factories S_1 , S_2 and S_3 with production capacity of 7, 9 and 18 units (in 100 s) per week of a product, respectively. These units are to be shipped to four warehouses D_1 , D_2 , D_3 and D_4 with requirements of 5, 8, 7 and 14 units (in 100's) per week, respectively. The transportation costs (in rupees) per unit between factories to ware houses are given in table below.

Warehouses Factories	Warehouses				Capacity
	D_1	D_2	D_3	D_4	
S_1	19	30	50	10	7
S_2	70	30	40	60	9
S_3	40	8	70	20	18
Demand	5	8	7	14	34

Obtain the initial basic solution using LC Method and also obtain the optimum solution using MODI method. (12)

b) Twenty-five successive observations on a stationary time series are given as follows:

30, 33, 32, 27, 25, 28, 29, 31, 35, 34, 38, 31, 23, 24, 34, 36, 29, 32, 38, 27, 22, 29, 20, 40, 37.

Calculate r_1, r_2, \dots, r_{10} and plot the correlogram.

(06)

TUTOR MARKED ASSIGNMENT

MSTL-001: Basic Statistics Lab

Course Code: MSTL-001

Assignment Code: MSTL-001/TMA/2021

Maximum Marks: 100

Note:

1. All questions are compulsory.
2. Solve the following questions in MS Excel 2007.
3. Take the screenshots of the final output/spreadsheet.
4. Paste all screenshots in the assignment booklet with all necessary hypotheses, interpretation, etc.

Q 1) A mobile manufacturing company distributed a newly launched mobile to 100 retail stores. These stores also sell another famous brand of mobile having same features. The manager of the company wants to compare the popularity of the newly launched mobile (say, Brand A) with the other popular mobile (say, Brand B). For this purpose, she selects a sample of 100 stores and noted the total number of sold mobiles of each brand. The data are recorded in the following table:

Store No.	Brand A	Brand B	Store No.	Brand A	Brand B
1	204	462	51	521	239
2	328	454	52	327	284
3	262	211	53	531	225
4	364	284	54	490	175
5	478	168	55	427	365
6	368	304	56	310	384
7	506	162	57	433	182
8	362	328	58	313	505
9	151	484	59	424	270
10	371	256	60	295	183
11	522	159	61	288	463
12	328	362	62	239	517
13	532	178	63	244	435
14	491	230	64	270	495
15	428	392	65	224	449
16	311	322	66	484	440
17	434	349	67	173	418
18	314	339	68	256	498
19	425	312	69	416	410
20	296	446	70	233	546
21	289	352	71	323	409
22	240	467	72	279	529
23	245	392	73	340	214
24	271	293	74	464	262
25	225	500	75	383	216
26	485	239	76	423	461
27	174	283	77	351	497

28	257	380	78	404	271
29	417	295	79	226	475
30	234	466	80	164	511
31	324	365	81	205	274
32	280	449	82	329	295
33	341	189	83	263	326
34	465	218	84	365	157
35	384	530	85	479	368
36	424	321	86	369	256
37	352	485	87	507	272
38	405	369	88	363	226
39	227	457	89	163	334
40	165	420	90	372	218
41	203	293	91	523	184
42	327	315	92	329	437
43	261	399	93	533	206
44	363	409	94	492	177
45	477	213	95	429	173
46	367	256	96	177	502
47	505	190	97	435	406
48	361	238	98	315	358
49	161	495	99	426	381
50	370	428	100	297	408

Answer the following:

- i) Which mobile brand has more average sales?
- ii) Which mobile brand shows greater variability in the sales?
- iii) Compute the skewness and kurtosis for the given data and interpret the results.
- iv) Determine the correlation between both mobiles.
- v) Compute suitable width of the class intervals for both brands,
- vi) Construct the continuous frequency distribution for both brands.
- vii) Represent the raw data as well as grouped data using suitable diagram/graph.

(2+4+4+3+2+4+6)

Q 2 (i) Twenty operators are working in a particular department of a call centre. The long-serving operators feel that they should have a promotion based on length of service built into their job structure. For assessment of their efficiency, the personnel department produces a score of efficiency based on various parameters. The efficiency scores along with their length of service are given below:

S. No.	Length of Service	Efficiency Scores
1	2	72
2	9	90
3	5	74
4	9	88
5	4	70

6	5	77
7	6	77
8	5	78
9	7	87
10	6	80
11	4	72
12	8	83
13	8	84
14	7	84
15	8	85
16	7	90
17	3	69
18	5	89
19	7	90
20	5	79

Compute the Spearman's rank correlation coefficient between the length of service and efficiency scores? Also compare this coefficient with Pearson's correlation coefficient.

- (ii) A scientist wants to check the yield of wheat from different varieties of wheat. In this regards, the yields of wheat from 16 plots, all of approximately equal fertility, when 4 varieties of wheat (Say, A, B, C and D) were cultivated are recorded in the following table:

Plot No.	1	2	3	4	5	6	7	8
Variety:	A	B	D	C	B	C	A	D
Yield:	42	44	39	41	43	44	44	36
Plot No.:	9	10	11	12	13	14	15	16
Variety:	B	D	A	C	B	A	B	C
Yield:	46	40	43	45	47	45	45	42

Assuming that the yields of wheat are normally distributed in each variety with approximately equal variances, test the hypothesis that the average yields of wheat in all four varieties of wheat are the same at 1% level of significance.

(12+13)

- Q 3** For the data given in Question 1, the manager also wishes to compare both brands to get the answers of the following questions:

- i) Is there enough evidence that the average sales of mobile of Brand A is more than the average sales of the Brand B at 5 % level of significance?
- ii) Are the variances of the distributions of mobiles of Brand A and Brand B equal at 5 % level of significance?

(13+12)

- Q 4** Suppose that a customer service manager of a mall wants to evaluate the service of the eight food counters in the mall. She/he hires seven evaluators with varied experience in food-service evaluation to act as raters. To reduce the effect of the variability from rater to rater, a randomised block design is applied considering raters serving as the blocks. The seven raters evaluate the service of each of the eight food counters in a random

order. A rating scale from 0 (low) to 100 (high) is used. The following table summarises the results:

	Food Counters							
Raters	A	B	C	D	E	F	G	H
1	70	61	82	74	68	59	80	81
2	77	75	88	76	75	73	86	87
3	76	67	90	80	74	65	82	88
4	80	63	87	76	78	61	85	86
5	84	66	92	84	82	64	90	91
6	78	68	94	86	76	66	80	92
7	77	75	88	76	75	73	86	84

The effect of evaluation of each rater on the service of food counters is normally distributed with approximately equal variances.

- i) Analyse the design at 2% level of significance.
- ii) Is the average service of the eight food counters significantly different? If the difference between the averages services of the eight restaurants is significant, do the pair-wise comparison between them.

(25)

TUTOR MARKED ASSIGNMENT

MSTL-002: Industrial Statistics Lab

Course Code: MSTL-002

Assignment Code: MSTL-002/TMA/2021

Maximum Marks: 100

Note:

1. All questions are compulsory.
2. Solve the following questions in MS Excel 2007.
3. Take the screenshots of the final output/spreadsheet.
4. Paste all screenshots in the assignment booklet with all necessary hypotheses, interpretation, etc.

Q 1 The marketing manager of a transportation network company offering taxi services in a metro city wanted to study the waiting times of customers to get a taxi during the peak hours. A subgroup of 15 customers was selected (one at each ten minutes interval during the hour) and the time in minutes was measured from the point each customer booked the taxi to when he or she began the trip. The results of 40 days period were as under.

Sample No.	1	2	3	4	5	6	7	8
Obs. 1	8.7	6.6	6.7	5.6	6.9	8.2	7.3	5.7
Obs. 2	6.4	6.3	7.5	8.7	7.1	6.7	8.8	8.1
Obs. 3	8.8	7.3	9.5	7.1	10	7.2	7.1	5.2
Obs. 4	6.1	6.7	8.1	9.1	7.5	7.1	8.7	8.5
Obs. 5	8.6	8.5	9.3	6.9	9.8	7.1	8.9	7
Obs. 6	5.6	5.5	7.5	7.9	6.3	6.9	8	7.3
Obs. 7	8	7.9	8.7	6.3	9.2	6.5	8.3	4.4
Obs. 8	5.3	5.9	7.3	8.3	6.7	8	7.9	7.7
Obs. 9	5.8	5.8	7.7	8.1	6.5	7.7	8.2	7.5
Obs. 10	8.2	8.1	8.9	6.5	9.4	6.7	8.5	6.3
Obs. 11	5.5	6.1	7.5	8.5	6.9	6.5	8.1	7.9
Obs. 12	8.1	7.9	8.8	6.3	9.2	8.6	8.3	6.4
Obs. 13	6.5	6.4	8.7	9.1	7.3	6.8	8.9	8.4
Obs. 14	9.3	9.2	10.1	7.3	10.7	7.5	8.9	5
Obs. 15	6.1	6.8	8.4	9.6	7.7	7.3	9.2	8.9

Sample No.	9	10	11	12	13	14	15	16
Obs. 1	9.3	9	9	5.8	7.1	6	8.5	6.7
Obs. 2	8.5	8.2	7.2	7.9	7.7	7.3	10.7	8
Obs. 3	7.5	5.4	7.2	8.5	9.1	8	6.4	6.5
Obs. 4	8.7	8.6	7.7	8.3	8.5	7.7	11.1	8.4
Obs. 5	8.1	7.3	7	8.3	7.5	7.9	6.2	6.3
Obs. 6	7.8	7.4	6.5	7.1	8.6	6.5	9.9	7.2
Obs. 7	8.4	6.7	6.4	7.7	7.8	7.2	5.6	5.7
Obs. 8	7.3	7.8	6.9	7.5	8.1	6.9	10.3	7.6

Obs. 9	8.1	7.6	6.7	7.3	8.8	6.7	10.1	7.4
Obs. 10	8.6	4.8	6.6	8	8	7.5	5.8	5.9
Obs. 11	8.2	8	7.1	7.8	8.3	7.1	10.5	7.8
Obs. 12	7.5	4.6	6.4	7.8	8.8	7.3	5.6	5.8
Obs. 13	9.1	8.6	7.5	8.2	10	7.5	11.5	8.3
Obs. 14	9.7	7.7	7.4	9	9	8.4	6.5	6.6
Obs. 15	8.4	9.1	7.9	8.7	7.8	8	11.9	8.8

Sample No.	17	18	19	20	21	22	23	24
Obs. 1	8.6	6.3	6.6	5.5	6.8	8.1	7.2	5.6
Obs. 2	5.6	6.9	8.5	8.6	7	6.6	8.7	8
Obs. 3	8.7	8.6	7.5	7	9.9	7.1	9	5.1
Obs. 4	6	5.2	8.6	9	7.4	7	7.9	8.4
Obs. 5	8.5	8.4	6.8	6.8	9.7	7	8.8	6.9
Obs. 6	6.6	5.4	7.6	7.8	6.2	8	7.9	7.2
Obs. 7	7.9	7.8	8.6	6.2	9.1	6.4	8.2	4.3
Obs. 8	5.2	5.8	7.8	8.2	6.6	6.2	7.1	7.6
Obs. 9	5	6.4	7.8	8	6.4	6	8.1	7.4
Obs. 10	8.1	8	8.8	6.4	9.3	6.6	8.4	4.5
Obs. 11	5.4	4.6	8	8.4	6.8	6.4	7.4	7.8
Obs. 12	8	7.8	8.7	6.2	9.1	6.4	8.2	6.3
Obs. 13	7.6	6.3	8.8	9	7.1	6.7	9.1	8.3
Obs. 14	9.2	9	10	7.1	10.5	7.3	9.5	8.3
Obs. 15	6	6.7	9	9.5	7.6	7.2	8.3	8.8

Sample No.	25	26	27	28	29	30	31	32
Obs. 1	7.5	9.2	8.9	5.7	7	8.5	8.4	6.6
Obs. 2	8.1	8.1	7.1	7.8	9.3	7.2	10.6	7.9
Obs. 3	9.1	6.7	7.1	8.4	7.4	7.9	5.2	6.4
Obs. 4	7.9	8.5	7.6	8.2	8.1	7.6	11.7	8.3
Obs. 5	8.9	5.9	6.9	8.2	8.4	7.8	6.1	6.2
Obs. 6	7.3	7.3	6.4	7	8.5	6.4	9.8	7.1
Obs. 7	8.3	5.9	6.3	7.6	8.7	7.1	4.4	5.6
Obs. 8	7.2	7.7	6.8	7.4	7.3	6.8	10.9	7.5
Obs. 9	7.5	7.5	6.6	7.2	8.7	6.6	10	7.3
Obs. 10	8.5	6.2	6.5	7.9	8.9	7.4	4.6	5.8
Obs. 11	7.4	7.9	7	7.7	7.5	7	11.1	7.7
Obs. 12	8.3	5.3	6.3	7.7	7.9	7.2	5.5	5.7
Obs. 13	8.4	8.5	7.3	8.1	9.9	7.4	11.3	8.2
Obs. 14	9.6	6.9	7.2	8.8	8.9	8.3	5	6.5
Obs. 15	8.3	8.9	7.8	8.6	8.4	7.9	12.6	8.7

Sample No.	33	34	35	36	37	38	39	40
------------	----	----	----	----	----	----	----	----

Obs. 1	8.5	5.2	6.5	9.7	6.2	7.6	6.5	9.1
Obs. 2	8.1	5.7	7.4	7.8	8.5	8.3	7.8	11.5
Obs. 3	8.6	8.5	9.3	7.7	9.1	6.8	8.6	6.8
Obs. 4	5.9	6.6	7.1	8.2	8.9	9.1	8.3	11.9
Obs. 5	8.4	8.3	9.1	7.5	9	8	8.4	6.6
Obs. 6	7.3	4.9	7.3	6.9	7.6	6.2	7	10.6
Obs. 7	7.8	7.7	8.5	6.8	8.3	8.3	7.8	6
Obs. 8	5.1	5.8	6.4	7.4	8.1	8.7	7.4	11.1
Obs. 9	7.5	5.2	9.8	7.1	7.9	8.8	7.2	10.8
Obs. 10	8	7.9	8.7	7.1	8.5	8.6	8	6.2
Obs. 11	5.3	6.1	6.6	7.6	8.3	8.9	7.6	11.3
Obs. 12	7.9	7.7	8.6	6.9	8.3	9.3	7.8	6
Obs. 13	8.4	8	8.4	6.6	7.4	9.2	10.4	8.4
Obs. 14	9.1	8.9	9.9	7.3	7.2	9.7	10.2	8.1
Obs. 15	5.9	6.7	7.3	10.4	10.2	7.3	8.2	11.9

The manager of this company needs to construct the suitable control charts for variability as well as average to infer whether the waiting times of customers for getting a taxi is under control or not. If it is out-of-control, also construct the revised control charts.

(25)

Q 2 A publisher recorded the total number of pages in 25 published books and also the number of typing errors that have been made in preparing the final print of the books.

The results are given in the following table:

Day	Number of pages	Errors
1	180	5
2	187	11
3	205	11
4	180	25
5	180	11
6	172	5
7	183	11
8	194	16
9	187	14
10	180	4
11	198	16
12	216	8
13	198	16
14	198	14
15	187	5
16	172	11
17	180	8
18	201	16
19	187	5
20	190	11
21	180	8

22	198	6
23	180	14
24	180	8
25	169	11
26	150	7
27	156	14
28	150	14
29	150	17
30	144	14
31	162	7
32	156	14
33	150	16
34	180	18
35	156	4
36	144	21
37	150	9
38	164	15
39	156	16
40	150	5

The publisher needs to set up a suitable control chart for the number of errors to check whether the number of errors in a state of control or not. Also computes the revised control limits, if necessary. (25)

Q 3 A researcher is interested to check the relationship between the salaries of workers involved in the production process of a company. To accomplish this, she/he has decided to develop a multiple regression model to predict their weekly salaries. For this purpose, he/she has selected a random sample of 50 workers involved in the production process. The information on their current monthly salaries in hundreds (Y), lengths of employment in months (X_1), and job classifications (X_2 ; 0 for technical job and 1 for clerical job) are summarised in the following table:

Employee	Y	X_1	X_2
1	495	74	1
2	406	51	0
3	567	130	1
4	523	25	1
5	575	178	1
6	437	42	0
7	664	242	1
8	491	57	1
9	472	72	0
10	407	129	1
11	378	17	1
12	725	318	1
13	600	296	0

14	440	39	0
15	662	280	1
16	523	116	1
17	428	19	1
18	535	94	1
19	533	193	0
20	528	49	1
21	446	26	0
22	528	40	1
23	498	51	1
24	478	48	1
25	507	32	0
26	478	24	1
27	645	234	1
28	577	281	0
29	554	335	1
30	654	336	1
31	566	77	1
32	433	90	0
33	466	89	1
34	365	30	0
35	677	225	0
36	473	36	1
37	644	305	1
38	685	316	1
39	356	11	0
40	444	23	1
41	478	94	0
42	408	81	0
43	456	58	0
44	409	22	0
45	691	359	0
46	463	69	0
47	436	93	0
48	413	16	1
49	734	412	0
50	463	69	0

- i) Prepare a scatter plot to get an idea about the relationship among the variables.
- ii) Fit a linear regression model and its related analysis at 1% level of significance.
- iii) Does the fitted regression model satisfy the linearity and normality assumptions?
- iv) Also, draw both fitted regression lines on the scatter plot.

(5+5+10+5)

- Q 4** The marketing manager of a transportation network company offering taxi services in a metro city wishes to improve customer service and taxi scheduling based on the daily levels of customers in the past 10 weeks. The numbers of customers during that period are given below:

Week	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	334	499	262	232	435	351	223
2	170	249	203	268	329	168	293
3	110	179	240	114	266	99	90
4	155	234	81	253	314	93	278
5	95	164	225	99	264	107	283
6	270	183	308	125	194	255	129
7	223	132	239	268	143	323	332
8	212	369	390	369	218	495	320
9	479	540	414	590	390	460	663
10	549	739	734	684	642	866	832

- i) Determine the seasonal indices for the given data using a 7-day moving averages.
- ii) Obtain the deseasonalised values.
- iii) Fit the appropriate trend for the deseasonalised data using the least-squares method by matrix approach that best describes the data.
- iv) Project the number of customers on Wednesday of the 52th week.
- v) Plot the original data, the deseasonalised data, and the trend values.

(8+3+7+3+4)