K.S.R. COLLEGE OF ENGINEERING (Autonomous) R 2018

SEMESTER - I

		L	T	P	C
MA18137	QUANTITATIVE TECHNIQUES FOR BUSINESS (Master of Business Administration)	3	1	0	4

Objective(s):

- *The topics enable the students to learn the basic concepts of probability and its distributions.*
- *To understand the concepts and applications of analysis of variance in business, to interpret samples and obtain the inferences by using the testing of hypothesis methods.*
- *To apply non–parameter test in business decision making to analyze the interpolation techniques, correlation and regression in business applications.*

UNIT - I INTRODUCTION TO STATISTICS & PROBABILITY [12]

Statistics – Basic definitions and examples of organizing statistical survey – Definition and problems - Measurement of central tendency and skewness – Concept of Probability - Basic definitions and rules for probability, conditional probability, Baye's theorem - problems. Definition - Probability Distributions : Poisson and Normal distributions.(Excluding proof).

UNIT - II ANALYSIS OF VARIANCE (ANOVA) [12]

Definition and concept of ANOVA - one way and Two way analysis of variance- Concept of Randomized block design – Design of experiments – Latin Square design and its applications.

UNIT - III TESTING OF HYPOTHESIS [12]

Definition - testing of hypothesis - Basic concepts in Testing of Hypothesis. - Testing significance for attributes (single and Two proportion test) – Testing of large samples 'z' - test – Testing of small samples 't'-test - confidence limit – F - test for two sample standard deviations - problems based on its applications.

UNIT - IV NON-PARAMETRIC METHODS [12]

Definition and concepts of non-parametric tests - Chi-square tests for independence of attributes and goodness of fit - Sign test for paired data, Rank sum test, Mann – Whitney U test and Kruskal Wallis H-test – problems based on its applications.

UNIT - V INTERPOLATION, CORRELATION & REGRESSION ANALYSIS [12]

Definition - Interpolation – Newton's Gregory forward interpolation and backward interpolation method (for equal intervals) – Lagrange's interpolation method (for unequal intervals). Definition - Correlation and Regression Analysis – problems.

Total (L:45 T:15) = 60 Periods

Course Outcomes: On Completion of this course, the student will be able to

- *To understand the basic concepts of probability and its distributions.*
- *To understand the concepts of analysis of variance in business.*
- *To interpret samples and obtain the inferences by using the testing of hypothesis methods.*
- *To apply non–parameter test in business research.*
- *To analyze the interpolation techniques, correlation and regression in business applications.*

Reference Books :

1. S.P. Gupta , Statistical Methods, Sultan Chand & Sons 2014
2. Levin R.I. and Rubin D.S., Statistics for Management, Prentice Hall of India Pvt. Ltd., New Delhi, 7th edition 2012
3. Srivatsava TN, Shailaja Rego, Statistics for Management, Tata McGraw Hill, Second edition 2012.
4. Anand Sharma, Statistics for Management, Himalaya Publishing House, Second Revised edition, 2013

A.S
Verified

Chairman (BoS)
Prof.R.VEERASAMY
Head of the Department (S & H)
K.S.R. College of Engineering
Tiruchengode - 637 215.

K.S.R. COLLEGE OF ENGINEERING (AUTONOMOUS)
TIRUCHENGODE – 637 215
DEPARTMENT OF BUSINESS ADMINISTRATION

NAME : S. DHAVAMANI
CLASS : I MBA
SUBJECT : MA18137 - QUANTITATIVE TECHNIQUES FOR BUSINESS

UNIT- I - INTRODUCTION TO STATISTICS AND PROBABILITY

2 MARKS

1. What's the definition of Statistics?

Statistics are usually defined as:

1. A collection of numerical data that measure something.
2. The science of recording, organizing, analyzing and reporting quantitative information.

2. Define statistical methods.

A method of analyzing or representing statistical data; a procedure for calculating a statistic

3. What are the different components of statistics ?

There are four components as per Croxton& Cowden

1. Collection of Data.
2. Presentation of Data
3. Analysis of Data
4. Interpretation of Data

4. What is the need for Statistics?

Statistics gives us a technique to obtain, condense, analyze and relate numerical data. Statistical methods are of a supreme value in education and psychology.

5. What are the various modes of data collection?

- Telephone
- Mail
- Online surveys
- Personal survey
- Mall intercept survey

6. What is sampling?

"Sampling" basically means selecting people/objects from a "population" in order to test the population for something. For example, we might want to find out how people are going to vote at the next election. Obviously we can't ask everyone in the country, so we ask a sample.

7. What are the types of data collection?

Qualitative Data

- Nominal, Attributable or Categorical data
- Ordinal or Ranked data

Quantitative or Interval data

- Discrete data
- Continuous measurements

8. What is tabulation of data?

Tabulation refers to the systematic arrangement of the information in rows and columns. Rows are the horizontal arrangement. In simple words, tabulation is a layout of figures in rectangular form with appropriate headings to explain different rows and columns. The main purpose of the table is to simplify the presentation and to facilitate comparisons.

9. What is presentation of data?

Descriptive statistics can be illustrated in an understandable fashion by presenting them graphically using statistical and data presentation tools.

10. What are the forms of presentation of the data?

Grouped and ungrouped data may be presented as :

- Pie Charts
- Frequency Histograms
- Frequency Polygons
- Ogives
- Boxplots

11. What are the measures of summarizing data?

- Measures of Central tendency: Mean, median, mode
- Measures of Dispersion: Range, Variance, Standard Deviation

12. Define mean, medianand mode?

Mean: The *mean* value is what we typically call the "average." You calculate the mean by adding up all of the measurements in a group and then dividing by the number of measurements.

Median: Median is the middle most value in a series when arranged in ascending or descending order

Mode: The most repeated value in a series.

13. Which measure of central tendency is to be used?

The measure to be used differs in different contexts. If your results involve categories instead of continuous numbers, then the best measure of central tendency will probably be the most frequent outcome (the mode). On the other hand, sometimes it is an advantage to have a measure of central tendency that is less sensitive to changes in the extremes of the data.

14. Give the empirical relationship between mean, median and mode?

Mode = 3 Median – 2 Mean (for asymmetric distribution).

Mean = Median = Mode (for symmetric distribution).

15. Define Skewness.

In probability theory and statistics, **skewness** is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The **skewness** value can be positive or negative, or even undefined. The qualitative interpretation of the skew is complicated.

16. What is Probability?

Probability is a way of expressing knowledge or belief that an event will occur or has occurred.

17. What is a random experiment?

An experiment is said to be a random experiment, if it's out-come can't be predicted with certainty.

18. What is a sample space?

The set of all possible out-comes of an experiment is called the sample space. It is denoted by 'S' and its number of elements are n(s).

Example; In throwing a dice, the number that appears at top is any one of 1,2,3,4,5,6. So here:

S ={1,2,3,4,5,6} and n(s) = 6

Similarly in the case of a coin, S={Head,Tail} or {H,T} and n(s)=2.

19. What is an event? What are the different kinds of event?

Event: Every subset of a sample space is an event. It is denoted by 'E'.

Example: In throwing a dice S={1,2,3,4,5,6}, the appearance of an event number will be the event E={2,4,6}.Clearly E is a sub set of S.

Simple event: An event, consisting of a single sample point is called a simple event.

Example: In throwing a dice, S={1,2,3,4,5,6}, so each of {1},{2},{3},{4},{5} and {6} are simple events.

Compound event: A subset of the sample space, which has more than on element is called a mixed event.

Example: In throwing a dice, the event of appearing of odd numbers is a compound event, because E={1,3,5} which has ’3′ elements.

20. What is the definition of probability?

If ‘S’ be the sample space, then the probability of occurrence of an event ‘E’ is defined as:

P(E) = n(E)/N(S) = number of elements in ‘E’
number of elements in sample space ‘S’

21. If P(A/B) = 0.2 and P(B) = 0.4 then find P(A∩B).

W.K.T., $P(A/B) = \frac{P(A\cap B)}{P(B)}$, $\quad P(A \cap B) = P(A/B).P(B)$,

$$P(A \cap B) = (0.2)(0.4), \qquad P(A \cap B) = 0.08.$$

22. Find P(A/B), if P(A) = 0.6, P(B) = 0.7 and P(A∪B) = 0.8

W.K.T.,$(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$P(A \cap B) = P(A) + P(B) - P(A \cup B),$$

$$P(A \cap B) = 0.6 + 0.7 - 0.8 = 0.5,$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.5}{0.7} = 0.7143$$

23. State Baye’s theorem.

Let $A_1, A_2, A_3, \ldots\ldots A_n$ be a set of n mutually exclusive and collectively exhaustive events. If B is another event such that P(B) is not zero, then $P(A_1/B) = \frac{P(A/B_1)\ P(A_1)}{\sum_{i=1}^{k} P(A/B_1)\ P(A_1)}$.

24. Define Poisson distribution.

A discrete frequency distribution which gives the probability of a number of independent events occurring in a fixed time.

The Poisson Distribution is given by the function $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $\quad x = 0,1,2, \ldots$

Here λ is the parameter of the distribution and we write X ~ P (X ; λ).

The mean and variance of the Poisson distribution are both equal to the parameter λ.

25. **If X is a Poisson variate such that P[x = 2] = 9 P[x = 4]+90 P [X = 6], find mean.**

The probability distribution for the R. V ‘X’ is given by, $P(X = x) = e^{-\lambda}\lambda^x / x!$, x= 0,1,2,… and $\lambda > 0$.

Given that P[x = 2] = 9 P[x = 4]+90 P [X = 6]

That is, $\quad e^{-\lambda}\lambda^2 / 2! = 9\, e^{-\lambda}\lambda^4 / 4! = 90\, e^{-\lambda}\lambda^6 / 6!$

Dividing by $e^{-\lambda}\lambda^2$, we get $\lambda^4 + 3\lambda^2 - 4 = 0 \Rightarrow (\lambda^2 + 4)(\lambda^2 - 1) = 0 \Rightarrow \lambda = 1$ (since $\lambda > 0$)

Hence the mean is 1.

26. **Give an examples for Poisson distribution may be successfully employed**

The number of printing mistakes at each page of book. The number of suicides report in Erode.

27. **Define Normal distribution.**

A function that represents the distribution of many random variables as a symmetrical bell-shaped graph.

A continuous random variable X is said to follow the normal distribution if its probability density function is given by $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$.

28. **Write the pdf of normal distribution, mean, variance and MGF.**

Ar.v ‘X’ is said to follow normal distribution with mean μ and variance σ2, if its density function Is given by f(x) = $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$, $\sigma > 0$, $-\infty < \mu < \infty$.

29. **X is normally distributed and the mean of x is 12 and the S.D is 4. Find P(X ≥ 20).**
    Given $\mu = 12$, $\sigma = 4$.
    When X = 12, $Z = (X-\mu)/\ \sigma = (20-12)/4 = 2$.
    Therefore, $P(X \geq 20) = P(Z \geq 2) = 0.5 - P\ (0 \leq Z \leq 2) = 0.5 - 0.4772 = 0.0228$.
30. **Write the characteristics of normal distribution.**
    i) Normal distributions are symmetric around their mean.
    ii) The mean, median, and mode of a normal distribution are equal.
    iii) The area under the normal curve is equal to 1.0.
    iv) Normal distributions are denser in the center and less dense in the tails.
    v) Normal distributions are defined by two parameters, the mean (μ) and the standard deviation (σ).
    vi) 68% of the area of a normal distribution is within one standard deviation of the mean.
    vii) Approximately 95% of the area of a normal distribution is within two standard deviations of the mean.
31. Mean, Variance and third central moment of Poisson distribution are ---------- ***Answer: equal***
32. Poisson distribution is not a ----------- distribution. ***Answer: Symmetrical***
33. For a normal distribution, coefficient of skewness is ------------. ***Answer: zero.***
34. The graph of the normal distribution is ------------------. ***Answer: bell shaped***
35. The normal distribution is a ---------------------------- probability distribution. ***Answer: two parameter.***
36. The -------- of the normal
37. Distribution lies at the centre of normal curve. ***Answer: mean***

<u>12 MARKS</u>

1. Describe the various methods of diagrammatic and graphical representation of statistical data with examples,
2. Find mean and median to the following data.

Class	0 - 8	8 - 16	16 - 24	24 - 32	32 - 40	40 - 48
Frequency	8	7	16	24	15	7

3. A bag contains 5 balls and it is not known how many of them are white. Two balls are drawn at random from the bag and they are noted to be white. What is the chance that all the balls in the bag are white?
4. Compute the mean and variance of Poisson distribution that has a double mode at x = 1 and x = 2.
5. The number of monthly breakdown of a computer is a R.V having Poisson distribution with mean equal to 0.8. Find the probability that this computer will function for a month

i) without a breakdown, ii)with only one breakdown, iii) with at least one breakdown.

6. A man is equally likely to choose any of the three routes A, B, C from his house to the railway station. The probabilities of missing the train by the respective routes are 1/20, 1/10, 1/5. One day he missed the train. What is the probability that the route chosen was C?
7. Bring out the main characteristics of Normal distribution.
8. In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and Standard deviation of the distribution.
9. Find the mean and SD of normal distribution in which 7% of the items are below 35 and 89% are below 63.
10. In an examination, it is laid down that a student passes if he secures 30% or more marks. He is placed in first, second and third division according as he secures 60% or more marks, between 45% and 60% marks and marks between 30% and 45% respectively. If he secures 80% or more marks, he gets

distinction. It is noticed from the results that 10% of the students failed and 5% of them obtained distinction. Assuming normal distribution of marks, what percentage of students placed in the second division?

11. If the actual amount of instant coffee which a filling machine puts into '6 – ounce' jars is a R.V having a normal distribution with S.D = 0.05 ounce and if only 3% of the jars are to contain less than 6 ounces of coffee, what must be the mean fill of these jars?
12. Explain the importance of normal distribution. The mean and s.d.of the wages of 6000 workers engaged in a factory are Rs.1,200 and Rs.400 respectively. Assuming the distribution to be normally distributed estimate
    (i)percentage of workers getting wages above 1,600
    (ii)number of workers getting wages between Rs.600 and 900
    (iii) number of workers getting wages between Rs.1,100 and Rs.1,500
13. If X is a normal variable with mean 30 and S.D = 5, then find
    a) $P[26 \leq X \leq 40]$ b) $P[X \geq 45]$ c) $P[|X - 30| > 5]$.

UNIT-II - TESTING OF HYPOTHESIS

<u>2 MARKS</u>

1. What do you mean by population and sample?

A population consists of collection of individual units, which may be persons or experimental outcomes, whose characteristics are to be studied. A sample is a portion of the population that is studied to learn about the characteristics of the population.

2. Differentiate Parameter and Statistics.

Population constants are called as Parameter and it is always denoted by Greek letters. Sample constants are called as Statistics and it is always denoted by English alphabets.

3. What is a statistical hypothesis?

A **statistical hypothesis** is an assumption about a population parameter. This assumption may or may not be true.

4. What is hypothesis testing?

Statisticians follow a formal process to determine whether to reject a null hypothesis, based on sample data. This process is called **hypothesis testing.**

5. Define Null and Alternative hypothesis.

In attempting to reach decisions about population on the basis of sample observations, we make assumptions about population, which are not necessarily true, are called statistical hypothesis

6. What are the procedures of testing hypothesis?

Fix the null and alternative hypothesis 2. Fix the level of significance 3. Write down the test statistics 4. Calculations 5.Inferences

7. Define the steps of hypothesis testing?

Hypothesis testing consists of four steps.

i) State the hypotheses. This involves stating the null and alternative hypotheses. The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false.

ii) Formulate an analysis plan. The analysis plan describes how to use sample data to evaluate the null hypothesis. The evaluation often focuses around a single test statistic.

iii) Analyze sample data. Find the value of the test statistic (mean score, proportion, t-score, z-score, etc.) described in the analysis plan.

iv) Interpret results. Apply the decision rule described in the analysis plan. If the value of the test statistic is unlikely, based on the null hypothesis, reject the null hypothesis.

8. What do you meant by critical region?

The critical region of a test of statistical hypothesis is that region of the normal curve which corresponds to the rejection of null hypothesis.

9. Define one tail and two tail test.

When two tails of the sampling distribution of the normal curve are used, the relevant test is called two – tailed test. The alternative hypothesis $H_1: \mu_1 \neq \mu_2$ is taken in two –tailed test for $H_0: \mu_1 = \mu_2$.

When only one tail of the sampling distribution of the normal curve is used, the test is described as one-tail test.

10. Define standard error.

The standard deviation of the sampling distribution is called the standard error.

11. What are the assumptions and applications of t-test?

Applications of t – test

(i) To test the significance of a single mean

(ii) To test the significance of the difference between two sample means

(iii) To test the significance of the coefficient of correlation.

Assumptions of Students t – Test

(i) The parwent population from which the sample is drawn is normal.

(ii) The sample observations are independent, i.e., the sample is random.

(iii) The population standard deviation σ si unknown.

12. What are the types of errors in testing hypothesis?

Type I Error – rejection of null hypothesis when it is correct

Type II error – acceptance of null hypothesis when it is wrong.

13. What is the F-test?

The F-test is used for comparisons of the components of the total deviation. For example, in one-way, or single-factor ANOVA, statistical significance is tested for by comparing the F test statistic

$F = \frac{\text{variance between items}}{\text{variance within items}} F^* = \frac{\text{MSTR}}{\text{MSE}}$ where

$\text{MSTR} = \frac{\text{SSTR}}{I-1}$, I = number of treatmentsand

$\text{MSE} = \frac{\text{SSE}}{n_T - I}$, n_T = total number of casesto the F-distribution with $I - 1, n_T - I$ degrees of freedom. Using the F-distribution is a natural candidate because the test statistic is the quotient of two mean sums of squares which have a chi-square distribution.

14. What are decision errors?

Two types of errors can result from a hypothesis test.

- **Type I error**. A Type I error occurs when the researcher rejects a null hypothesis when it is true. The probability of committing a Type I error is called the **significance level**. This probability is also called **alpha**, and is often denoted by α.
- **Type II error**. A Type II error occurs when the researcher fails to reject a null hypothesis that is false. The probability of committing a Type II error is called **Beta**, and is often denoted by β. The probability of *not* committing a Type II error is called the **Power** of the test.

15. How to arrive at a decision on hypothesis?

The decision rules can be taken in two ways – with reference to a P-value or with reference to a region of acceptance.

- P-value. The strength of evidence in support of a null hypothesis is measured by the **P-value**. Suppose the test statistic is equal to *S*. The P-value is the probability of observing a test statistic as extreme as *S*, assuming the null hypotheis is true. If the P-value is less than the significance level, we reject the null hypothesis.
- Region of acceptance. The **region of acceptance** is a range of values. If the test statistic falls within the region of acceptance, the null hypothesis is not rejected. The region of acceptance is defined so that the chance of making a Type I error is equal to the significance level.The set of values outside the region of acceptance is called the **region of rejection**. If the test statistic falls within the region of rejection, the null hypothesis is rejected. In such cases, we say that the hypothesis has been rejected at the α level of significance.

12 MARKS

1. A sample of 900 members is found to have a mean 3.5cm. can it be reasonably regarded as a simple sample from a large population whose mean is 3.38cm and a standard deviation 2.4cm?
2. A manufacturer claims that his synthetic fishing line has a mean breaking strength of 8 kg and S.D. 0.5 kg. Can we accept his claim if a random sample of 50 lines yielded a mean breaking strength of 7.8kg? Use 1% level of significance.
3. The average hourly wage of a sample of 150 workers in plant A was Rs. 2.26 with a S.D. Rs. 1.08. The average wage of a sample of 200 workers in plant B was Rs. 2.87 with S.D. Rs. 1.28. Can an applicant safely assume that the hourly wage paid by plant B are higher those paid by plant A?
4. Given that on the average 4% of insured men of age 65 die within a year and that 60 of the particular group of 1000 such men died within a year. Can this group be regarded as the representative sample?
5. Manufacturers claim that only 4% of his product supplied by him is defective. A random sample of 600 products contained 36 defectives. Test the claim of the manufacturer.
6. In a large city A 20% of random sample of 900 school boys had defective eye-sight. In another large city B, 18.5% of a random sample of 1600 school boys had the same defect. Is the difference between the two proportions significant?
7. In a simple sample of 600 men from a large city, 400 are found to be car owners. In one of 900 from another large city, 450 arte owners. Do the data indicate that the cities are significantly different with respect to car owning among men?
8. For two samples, the following data are given.

$n_1 = 1000, \quad \bar{x}_1 = 67.42; \quad s_1 = 2.58$

$n_2 = 1200, \quad \bar{x}_2 = 67.25' \quad s_2 = 2.50$

Is the difference between the standard deviations significant?

9. The mean yield of two seats of plots and their variability are as given below. Examine whether the difference in the variability in yields is significant.

	Set of 40 plots	Set of 60 plots
Men yield per plot	1258	1243
S.D. per plot	34	28

10. Ten individuals are chosen at random from a population and their heights are found to be in inches 63,63,67,68,69,70,70,71,71. in the light of this data, discuss the suggestion that the mean height in the universe is 66 inches.

11. The average breaking strength of steel rods is specified to be 17.5lbs. to test this, a sample of 14 rods was tested and gave the following results (in units of 100 lbs):
15,18,16,27,17,17,15,17,20,19,17,18.Is the result of the experiment significant? Also obtain the 95% fiducial limits from the sample for the average breaking strength of steel rod.

12. Two independent samples from normal populations with equal variances gave the following results. Test for the equality of means.

Sample	Size	Mean	S.D.
1	16	23.4	2.5
2	12	24.9	2.8

13. In a test examination to two groups of students, the marks obtained were as follows

I group	18	20	36	50	49	36	34	49	41
II group	29	28	26	35	30	44	46		

14. Examine the significance of difference between the average marks obtained mark secured by the students of the above two groups.

15. The weight gains in ponds under two systems of feeding of calves of 10 pairs of identical twins is given below:

Weight gain under	1	2	3	4	5	6	7	8	9	10
System A	43	39	39	42	46	43	38	44	51	43
System B	37	35	34	41	39	37	35	40	48	36

Discuss whether the difference between the two systems of feeding is significant.

16. It is known that the mean diameters of rivers produced by two firms A and B are practically the same but the standard deviations may differ. For 22 rivets produced by A, the S.D. is 2.9m, while for 16+ rivets manufactured by B, the S.D is 3.8. Test whether the products of A have the same variability as those of B.

17. The data given below are the qualities of ten items (in proper units) produced by two processes A and B. Test whether the variability of quality may be taken to be the same for the two processes.

Process A:	3	7	5	6	5	4	4	5	3	3
Process B:	8	5	7	8	3	2	1	6	5	7

18. 200 digits were chosen at random from a set of tables. The frequencies of the digits were

Digit: 0 1 2 3 4 5 6 7 8 9

Frequency:	18	19	23	21	16	25	22	20	21	15

Use chi – square test to access the correctness of the hypothesis that the digits were distributed in equal number in the tables from which these were chosen.

19. The theory predicts the population of beans in the four groups A, B, C and D should be 9:3:3:1. In an experiment among 1600 beans, the number in the four groups was 882,313,287 and 118. Does the experimental result support the theory?

20. In a locality 100 persons were randomly selected and asked about their educational achievements. The results are given as below

	Education		
	Middle	High school	college
Male	10	15	25
Female	25	10	15

Can you say that education depends on gender?

21. Find out whether the new treatment is comparatively superior to the conventional one from the following data.

Favourable	non- favorable	total	
Conventional	40	70	110
New	60	30	90
Total	100	100	200

UNIT-III - NON-PARAMETRIC METHODS

<u>2 MARKS</u>

1. **Point out applications of chi-square test.**

   1. Chi-square test for goodness of fit 2. Chi-square test for independence of attributes 3. Chi-square test for Homogeneity

2. **Define chi – square distribution.**

   If X_i, (i = 1,2,…,n) are n independent normal variates with mean μ_i and variance σ_i^2 (i = 1,2,…,n) then

   $\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2$ is a chi – square variate with n degrees of freedom.

3. **Give the formula for Chi Sqare test of independence for**

a	**b**
c	**d**

   $$\chi^2 = \frac{N(ad - bc)^2}{(a+b)\,(c+d)\,(a+c)\,(b+d)}, \text{ where } N = a + b + c + d.$$

4. **Write down the probability density function of chi – square distribution.**

The probability density function of the χ^2- distribution is given by

$$f(\chi^2) = \frac{1}{2^{n/2}\sqrt{\frac{n}{2}}}(\chi^2)^{n/2-1}e^{-\chi^2/2}, \quad 0 < \chi^2 < \infty,$$ where n is the degrees of freedom.

5. **Write the uses of chi-square test.**
   - To test if the hypothetical value of the population variance is $\sigma^2 = \sigma_0^2$ (say).
   - To test the 'goodness of fit'. It is used to determine whether an actual sample distributions matches a known theoretical distribution.
   - To test the independence of attributes.
   - To test the homogeneity of independent estimates of the population correlation coefficient.

6. **Name any four non – parametric tests.**
   i) Sign test, ii) Rank sum test, iii) One sample run test, iv) Rank correlation test.

7. **What is sign test?**
   Sign test is a non-parametric statistical test for identifying differences between two populations based on the analysis of nominal data.
8. **What is Rank sum test.**
   Nonparametrical statistical test for identifying differences between two or more populations based on the analysis of two or more independent samples one from each population are used. Mann-Whitney 'U' test and Kruskal-Wallis test are called Rank-sum test because the test depends on the ranks of the sample observations.

9. **When is** Mann-Whitney 'U' test and Kruskal-Wallis test are used?
   Mann-Whitney test is used when there atre only two populatios whereas Kruskal-Wallis test is employed when more than two populations are involved.

10. **Write Mann-Whiteny U-test Statistic.**
    Mann-Whitney test A statistical test of the probability that two independent sets of observations come from the same population. The Mann-Whitney test is independent of distribution and can be used when the t test is inappropriate.
11. **Define Kruskal-Wallis test or H-test.**
    It is used to test whether two populations are identical. The hypothesis for H-test with k > 3 populations can be written as follows:
    Null hypothesis: $\mu_1 = \mu_2 = \mu_{3.}$ (all the populations are identical)
    Alternative hypothesis: $\mu_1 \neq \mu_2 \neq \mu_3$. (all the populations are not identical)

12. **Define Run test with an example.**
    A run is a subsequence of one or more identical symbols representing a common property of the data (or) A run is a sequence of identical elements that are preceded and followed by different elements or no element at all.
    Example: Suppose that 12 people have been selected to constitute a committee, let us denote the male by M and female by F. Arrange the people according to the sex say

MMFFFMFFMMMM such a grouping are called runs. Here the total of 5 runs.

12 MARKS

1. 200 digits were chosen at random. Their frequencies are given below. Test whether the digits were distributed equally.

Digits	0	1	2	3	4	5	6	7	8	9
Frequency	18	19	23	21	16	25	22	20	21	15

2. What are the merits of Non-parametric methods over parametric methods?
3. Explain the Sign test for comparing paired samples with appropriate examples.
4. A food inspector examined 15 jars of a certain brand of jam to determine the percent of foreign impurities. The following data were recorded.
   2.4 2.3 1.7 1.7 2.3 1.2 1.1
   3.6 3.1 1.0 4.2 1.6 2.5 2.4 2.3
   Test the hypothesis that the average percent of impurities in this brand of jam is 2.5%.
   Use $\alpha = 0.01$.
5. The following data represent the monthly sales (inRs.) of a certain retail store in a leap year. Examine if there is any seasonality in the sales. 6100, 5600, 6350, 6050, 6250, 6200, 6300, 6250, 5800, 6000, 6150 and 6150.
6. Mendal's theory predicts the proportion of beans in four groups A,B,C,D should be in the ratio 9:3:3:1. In an experiment, the numbers in the four groups are respectively 882, 313, 287 and 118. Does the experimental result support the theory at 5% and 2% levels?

7. The following table shows the distribution of digits chosen in random from a telephone directory.

Digits	0	1	2	3	4	5	6	7	8	9
Frequency	1026	1107	997	966	1075	933	1107	972	964	853

   Calculate chi – square statistics and test 5% level whether the digits may be taken to occur equally frequently in the directory.
8. A company's trainees are randomly assigned to groups which are taught a certain inspection procedure by three different methods. At the end of the period they are tested for inspection performance quality.
   Method A : 80 83 79 85 90 68
   Method B :82 84 60 72 86 67 91
   Method C : 93 65 77 78 88
   Use the H –test to determine whether the methods are equally effective at 5% level of significance.
9. In an experiment on immunization of cattle from tuberculosis the following results were obtained.

	Affected	Not affected
Inoculated	12	26
Not inoculated	16	6

   Calculate chi – square statistics and discuss the effect of vaccine in controlling susceptibility to Tuberculosis.
10. Explain Run test and Kruskal Wallis H – test.
11. The following data are a random sample of consumer's income and expenditure on certain luxury items. Compute the Spearman rank correlation coefficient and test for the existence in the population.
    Income : 23 17 34 56 49 31 28 80 65 40 26
    Luxury items : 10 50 120 225 90 60 55 340 170 25 80
12. The ranks of 10competitors in a musical test given by 3 judges (A,B,C) are respectively (1,3,6),(6,5,4),(5,8,9),(10,4,8),(3,7,1),(2,10,2),(4,2,3),(9,1,10),(7,6,5) and (8,9,7). Identify the pair of judges having the nearest approach to common likings in music.

UNIT-IV - CORRELATION, REGRRESSION AND ANNOVA

2 MARKS

1. **What is correlation?**
   Correlation is a measure of association between two variables. The variables are not designated as dependent or independent.

2. **What can be the values for correlation coefficient?**
   The value of a correlation coefficient can vary from -1 to +1. A -1 indicates a perfect negative correlation and a +1 indicated a perfect positive correlation. A correlation coefficient of zero means there is no relationship between the two variables.

3. **What is the interpretation of the correlation coefficient values?**
   When there is a negative correlation between two variables, as the value of one variable increases, the value of the other variable decreases, and vise versa. In other words, for a negative correlation, the variables work opposite each other. When there is a positive correlation between two variables, as the value of one variable increases, the value of the other variable also increases. The variables move together.

4. **What is simple regression?**
   Simple regression is used to examine the relationship between one dependent and one independent variable. After performing an analysis, the regression statistics can be used to predict the dependent variable when the independent variable is known. Regression goes beyond correlation by adding prediction capabilities.

5. **What do you mean by an experiment?**
   An experiment is a device of getting an answer to the problem under consideration.

6. **Define treatment**.
   Various objects of comparison in a comparative experiment are termed as treatment.

7. **What is ANOVA?**
   Analysis of variance (ANOVA) is a collection of statistical models and their associated procedures in which the observed variance is partitioned into components due to different sources of variation. ANOVA provides a statistical test of whether or not the means of several groups are all equal.
8. **Why is ANOVA helpful?**
   ANOVAs are helpful because they possess a certain advantage over a two-sample t-test. Doing multiple two-sample t-tests would result in a largely increased chance of committing a type I error. For this reason, ANOVAs are useful in comparing three or more means.
9. **What are the assumption in ANOVA?**
   The following assumptions are made to perform ANOVA:
   - Independence of cases – this is an assumption of the model that simplifies the statistical analysis.
   - Normality – the distributions of the residuals are normal.
   - Equality (or “homogeneity”) of variances, called homoscedasticity — the variance of data in groups should be the same. Model-based approaches usually assume that the variance is constant. The constant-variance property also appears in the randomization (design-based) analysis of randomized experiments,

where it is a necessary consequence of the randomized design and the assumption of *unit treatment additivity* (Hinkelmann and Kempthorne): If the responses of a randomized balanced experiment fail to have constant variance, then the assumption of *unit treatment additivity* is necessarily violated. It has been shown, however, that the F-test is robust to violations of this assumption.

8. What is the logic of ANOVA?

Partitioning of the sum of squares

The fundamental technique is a partitioning of the total sum of squares (abbreviated SS) into components related to the effects used in the model. For example, we show the model for a simplified ANOVA with one type of treatment at different levels.

$$SS_{\text{Total}} = SS_{\text{Error}} + SS_{\text{Treatments}}$$

So, the number of degrees of freedom (abbreviated df) can be partitioned in a similar way and specifies the chi-square distribution which describes the associated sums of squares.

$$\text{df}_{\text{Total}} = \text{df}_{\text{Error}} + \text{df}_{\text{Treatments}}$$

9. What are the applications of completely randomized design?

(i) It is most useful in laboratory techniques and methodological studies.

(ii) It is recommended in situations where an appreciable fraction of units is likely to be destroyed.

10. Give the ANOVA table for C.R.D.

Source of Variation	Degrees of freedom	Sum of square	Mean sum of square	Variance ratio
Treatments Errors	v-1 n-v	S_T^2 S_E^2	$s_T^2 = S_T^2/$ v-1 $s_E^2 = S_E^2/$ n-v	$F_T = s_T^2/ s_E^2$
Total	n-1	$S_T^2 + S_E^2$		

11. What are the advantages of R.B.D.?

1. Accuracy 2. Flexibility 3. Ease of analysis

12. Give the ANOVA table for R.B.D.

Source of Variation	Degrees of freedom	Sum of square	Mean sum of square	Variance ratio
Treatments Blocks Errors	t-1 r-1 (t-1)(r-1)	S_T^2 S_B^2 S_E^2	$s_T^2 = S_T^2/$ t-1 $s_B^2 = S_B^2/$ r-1 $s_E^2 = S_E^2/$ (t-1)(r-1)	$F_T = s_T^2/ s_E^2$ $F_B = s_B^2/ s_E^2$
Total	rt-1	$S_T^2 + S_B^2 + S_E^2$		

13. Write any two differences between RBD and LSD.

S.No.	LSD	RBD
1.	The number of treatment is equal to the number of replications.	There is no such restrictions on treatments and replications.

2.	In the field layout, LSD can be performed on a square field.	In the field layout, RBD can be performed either on a square or rectangular field.
3.	It controls the variations between the rows and columns.	It controls the effect of one direction either row or column.
4.	LSD is known to be suitable for a case when the number of treatments is between 5 and 12 since the square becomes large and does not remain homogeneous.	RBD can be used for any number of treatments.

14. What are the principles of design of experiment?

1. Replication 2. Randomization 3. Local control

15. What do you mean by replication, randomization and local control?

Replication: replication means repetition of treatment under investigation.

Randomisation: a processes of assigning treatment to various experimental unit in a purely chance manner

Local control: The process of reducing the experimental error by dividing the relatively heterogeneous experimental area into homogeneous blocks is known as local control.

16. What are the advantages of Latin Square design?

i. It control the variations more than CRD and RBD
ii. LSD is an incomplete 3 way layout
iii. Very simple statistical analysis
iv. More than one factor can be investigated.

17. What do you mean by factorial experiment?

The effects of several factors of variation studied and investigated simultaneously when the treatments being app the combinations of different factors under study is called factorial experiment.

18. Name three important designs of experiments.

a. Completely randomized design (C. R. design).
b. Randomized block design (R. B. design).
c. Latin square design (L. S. design).

19. Compare CRD and RBD.

i) RBD is more efficient than CRD for most types of experimental work.
ii) In CRD, grouping of the experimental size so as to allocate the treatments at random to the experimental units is not done. But in RBD, treatments are allocated at random with in the units of each stratum.
iii) RBD is more flexible than CRD since no restriction are placed on the number of treatments or the number of replications.

12 MARKS

1. Obtain regression equation of Y on X and estimate Y when X=55 from the following

X	40	50	38	60	65	50	35
Y	38	60	55	70	60	48	30

2. Calculate Spearman's rank correlation for the following data.

Rank of X	2	4	3	5	1
Rank of Y	1	3	4	5	2

3. Find the correlation coefficient for the following data.

X	10	14	18	22	26	30
Y	18	12	24	6	30	36

4. Obtain the equations of lines of regression for the following data and hence estimate Y for X=75 and X for Y=70.

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

5. Three process A, B and C is tested to see whether the inputs are equivalent. The following observations of output are made;

A:	10	12	13	11	10	14	15	13
B:	9	11	10	12	13			
C:	11	10	15	14	14	13		

Analysis the experiment and give your conclusion.

6. Three varieties of coal were analyzed by four chemists and the ash content in the varieties was found to be as under

Varieties	1	2	3	4
A	8	5	5	7
B	7	6	4	4
C	3	6	5	4

Do the varieties differ significantly in their ash content?

7. An experiment was designed to study the performance of 4 different detergents for cleaning fuel injectors. The following "cleanness " readings were obtained with specially designed equipment for 12 tanks of gas distributed over 3 different models of engines.

	Engine 1	Engine 2	Engine 3
Detergent A	45	43	51
Detergent B	47	46	52
Detergent C	48	50	55
Detergent D	42	37	49

Test whether there are differences in the detergents or in the engines at 1% level of significance.

8. Carry out an analysis of variance to the following data related to sales of an item in different seasons.

		Sales men			
		A	B	C	D
Seasons	Summer	36	36	21	35
	Winter	28	29	31	32
	Monsoon	26	28	29	29

9. A varietal trail was conducted at a research station. The design adopted for the same was five randomized block of 6 plots each. The yield in lb. per plot obtained from the experiment as follows.

Blocks /Varieties	1	2	3	4	5	6
I	30	23	34	25	20	13
II	39	22	28	25	28	32
III	56	43	43	31	49	17
IV	38	45	36	35	32	20
V	44	51	23	58	40	30

Analyze the design and comment on your findings.

10. Perform an analysis of variance to the following LSD showing the effects of the fertilizers A,B,C,D on the yield of wheat.

A18	C21	D25	B11
D22	B12	A15	C19
B15	A20	C23	D24
C22	D21	B10	A17

11. Set up the analysis of variance for the following results of a Latin square design.

A(12)	C(19)	B(10)	D(8)
C(18)	B(12)	D(6)	A(7)
B(22)	D(10)	A(5)	C(21)
D(12)	A(7)	C(27)	B(17)

12. The following are the results of the latin square experiment on the effects of five manorial treatment A, B, C, D and E on the yield of sugarcane. Test whether the treatments are equally effect?

B(405)	A(525)	E(463)	D(441)	C (481)
C(325)	D(445)	B(429)	A(513)	E(493)
E(471)	B(492)	A(472)	C(381)	D(410)
A(552)	C(431)	D(425)	E (572)	B(451)
D(430)	E(469)	C(432)	B(467)	A(460)

13. Analyze the variance in the following Latin square of yields (in Kgs) of paddy where A,B,C,D denote the different methods of cultivation

D122	A121	C123	B122
B124	C123	A122	D125
A120	B119	D120	C121
C122	D123	B121	A122

Examine whether the different methods of cultivation have given significantly different yields

14. In order to determine whether there is significant difference in the durability of 3 makes of computers, sample of size 5 are selected from each make and the frequency of repair during the first year of purchase is observed the results are as follows:

A	5	6	8	9	7
B	8	10	11	12	4
C	7	3	5	4	1

In view of the above data, what conclusion can you draw?

15. Three different machines are used for a production. On the basis of the outputs, set up One-way ANOVA table and test whether the machines are equally effective.

OUTPUTS		
Machine I	Machine II	Machine III
10	9	20
15	7	16
11	5	10
10	6	14

UNIT-V - INTERPOLATIONM AND EXTRAPOLATION

2 MARKS

1. **Define Interpolation and Extrapolation?.**
   The process of finding the value of a function inside the given range is called Interpolation.

2. **Why the polynomial interpolation is preferred mostly?.**
   a. They are simple forms of functions which can be easily manipulated.
   b. Computations for definite values of the argument, integration and differentiation of such Functions are easy.
   c. Polynomials are free from singularities where as rational functions or other types, do have Singularities.

3. **State True or False:-**"Newton's forward interpolation formula is applicable only if the interval of differencing 'h' is constant"**.**Answer : True.

4. **Write the Newton's forward interpolation formula. Give the meaning of first two terms and first three terms of it.Y(x)** = y_o+
   $\frac{u}{1!}\Delta y_o + \frac{u(u-1)}{2!}\Delta y_o + \frac{u(u-1)(u-2)}{3!}\Delta^2 y_o + \ldots + \frac{u(u-1)\ldots(u-(n-1))}{n!}\Delta^n y_o$ .Where u = $\frac{x-x_o}{h}$. In the

   above formula, the first two terms will give the linear interpolation and the first three terms will give a parabolic interpolation.

5. **Write Newton's Backward Interpolation formula.**
   Y(x) =y_n + $\frac{v}{1!}\nabla y_n + \frac{v(v+1)}{2!}\nabla^2 y_n + \frac{v(v+1)(v+2)}{3!}\nabla^2 y_n + \ldots + \frac{v(v+1)\ldots(v+(n-1))}{n!}\nabla^n y_n$.

   Where v = $\frac{x-x_n}{h}$.

6. **State True or False :-**" The Newtons backward interpolation formula involves the backward difference operator, it is named as backward interpolation formula". Ans : True.

7. **State Newton's divided difference formula.**
   $f(x)= f(x_o) + (x-x_o)f(x_o,x_1) + (x-x_o)(x-x_1)\, f(x_o,x_1,x_2) + \ldots + (x-x_o)(x-x_1)\ldots(x-x_{n-1})\, f(x_o,x_1,x_2\ldots x_n)$.

8. State the error in Newton's forward interpolation and Newton's backward interpolation formula.

Error in forward interpolation

$$f(x) - p_n(x) = \frac{u(u-1)(u-2)...(u-n)}{(n+1)!} h^{n+1} f^{n+1}(c), where u = \frac{x-x_o}{h}.$$

Error in backward interpolation

$$f(x) - p_n(x) = \frac{v(v+1)(v+2)...(v+n)}{(n+1)!} h^{n+1} y^{n+1}(c), where \quad v = \frac{x-x_n}{h}.$$

9. State True or False:-" Eighth order difference of a polynomial of degree five is always zero". Ans : True

10. Fill in the blanks:-The nth order difference of the polynomial $p = a_o x^n + a_1 x^{n-1} + \ldots + a_n$ is <u>Constant</u>.

11. Say True or False:-"Lagrange's interpolation formula cannot be used when the base points are equally spaced". Ans : False

12. Newton's divided difference formula for equal interval is called <u>Newton's Gregory forward difference formula</u>.

13. Say True or False:-"Lagrange's interpolation formula can be used whether the arguments are equally spaced or not". Ans : True

14. Say True or False:-" The nth differences of a polynomial of degree n are zeros". Ans : True

15. State Lagrange's Interpolation formula.

Let y = f(x) be a function such that f(x) takes the values $y_o, y_1, \ldots, y_n$ corresponding to $x_o, x_1, \ldots x_n$, then

$$f(x) = \frac{(x-x_1)(x-x_2)...(x-x_n)}{(x_o-x_1)(x_o-x_2)...(x_o-x_n)} y_o + \frac{(x-x_0)(x-x_2)...(x-x_n)}{(x_1-x_0)(x_1-x_2)...(x_1-x_n)} y_1 + ... + \frac{(x-x_o)(x-x_1)...(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)...(x_n-x_{n-1})} y_n$$

16. State inverse Lagrange's Interpolation formula.

$$x = \frac{(y-y_1)(y-y_2)...(y-y_n)}{(y_o-y_1)(y_o-y_2)...(y_o-y_n)} x_o + \frac{(y-y_0)(y-y_2)...(y-y_n)}{(y_1-y_0)(y_1-y_2)...(y_1-y_n)} x_1 + ... + \frac{(y-y_o)(y-y_1)...(y-y_{n-1})}{(y_n-y_0)(y_n-y_1)...(y_n-y_{n-1})} x_n$$

17. When will you use Newton's backward interpolation formula.

We can apply the Newton's backward interpolation if the unknown value lies near the end of the table value.

18. Using Lagrange's interpolation formula, find the polynomial for

x	**0**	**1**	**3**	**4**
y	**-12**	**0**	**0**	**12**

y = f(x) =

$$\frac{(x-x_1)(x-x_2)...(x-x_n)}{(x_o-x_1)(x_o-x_2)...(x_o-x_n)} y_o + \frac{(x-x_0)(x-x_2)...(x-x_n)}{(x_1-x_0)(x_1-x_2)...(x_1-x_n)} y_1 + ... + \frac{(x-x_o)(x-x_1)...(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)...(x_n-x_{n-1})} y_n$$

$$y = \frac{(x-1)(x-3)(x-4)}{(0-1)(0-3)...(0-4)}(-12) + \frac{(x)(x-3)(x-4)}{(1-0)(1-2)(-4)} 0 + \frac{x(x-1)(x-4)}{(3-0)(3-1)(3-4)} 0 + \frac{x(x-1)(x-3)}{(4-0)(4-1)(4-3)} x$$

(i.e.) $y(x) = 2x^3 - 12x^2 + 22x - 12$.

19. Obtain the interpolation quadratic polynomial for the given data by using Newton's forward difference formula.

x	0	2	4	6
y	-3	5	21	45

Solution : The finite difference table is

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
0	-3			
2	5	8		
4	21	16	8	
6	45	24	8	0

$u = \frac{x-0}{2} = \frac{x}{2}$. Therefore $f(x) = y_o + u\,\Delta y_o + \frac{u(u-1)}{2!}\Delta^2 y_o = x^2 + 2x - 3.$

20. Find the parabola of the form $y = ax^2+bx+c$ passing through the points (0,0) (1,1) and (2,20).

By Lagrange's formula $y = \frac{(x-1)(x-2)}{(0-1)(0-2)}(0) + \frac{(x-0)(x-2)}{(1-0)(1-2)}1 + \frac{(x-0)(x-1)}{(2-0)(2-1)}20 = 9x^2 - 8x.$

21. Write the Lagrange's fundamental polynomial $L_o(x)$ and $L_1(x)$ that satisfy the condition $L_0(x)+L_1(x)$ for the data $[x_o,f(x_o)]$, $[x_1,f(x_1)]$.

Soln. : $L_o(x) = \frac{x-x_1}{x_o-x_1}$ *and* $L_1(x) = \frac{x-x_o}{x_1-x_o}$.

22. Say True or False.

"Lagrange's interpolation formula is applicable for both equal and unequal intervals". Ans:True.

23. Find the missing value of the table given below. What assumption have you made to find it?.

Year	**:**	**1917**	**1918**	**1919**	**1920**	**1921**
Export (in tons)	**:**	**443**	**384**	**--**	**397**	**467**

Sol.: Since four values are given, we assume that we have a third degree polynomial and hence its fourth order differences of $P_3(x)$ are zeros.

Let $u_0 = 443$, $u_1 = 384$, $u_2 =?$, $u_3 = 397$, $u_4 = 467$.

Clearly $\Delta^4 u_o = 0 \Rightarrow (E-1)^4 u_0 = 0 \Rightarrow (E^4 - 4E^3 + 6E^2 - 4E + 1)u_0 = 0$

$\Rightarrow u_4 - 4u_3 + 6u_2 - 4u_1 + u_o = 0 \Rightarrow 467 - 4(397) + 6u_2 - 4(384) + 443 = 0.$

$\Rightarrow 6u_2 = 2214 \Rightarrow u_2 = 369.$

24. Find the missing value of the following table.

X : 0	**1**	**2**	**3**	**4**
Y: 1	**2**	**4**	**-**	**16**

Explain why y(x=3) is not $2^3 = 8$ in your answer.

Soln. : Since only four values are given , assume that we have a third degree polynomial and hence it fourth order difference is zero.

Let $y_o= 1$, $y_1 = 2$, $y_2 = 4$, $y_3 = ?$ and $y_4 = 16$.

Clearly $\Delta^4 yo = 0 \Rightarrow (E-1)^4 y_0 = 0 \Rightarrow (E^4 - 4E^3 + 6E^2 - 4E + 1)\, y_0 = 0$

$\Rightarrow y_4 - 4y_3 + 6y_2 - 4y_1 + y_o = 0 \Rightarrow 16 - 4y_3 + 24 - 8 + 1 = 0 \Rightarrow 4y_3 = 33 \Rightarrow y_3 = 8.25.$

By looking at the table, we guess $y = 2^x$ is the function from which the table is created. So$y(3) = 2^3 = 8$.

12 MARKS

1. Using Lagranges interpolation formula find y(10) given that y(5) = 12, y(6) = 13, y(9) = 14 and y(11) = 16.

2. Find the missing term in the following table

x :	0	1	2	3	4
y :	1	3	9	-	81

3. From the data given below find the number of students whose weight is between 60 to 70.

Wt (x) :	0-40	40-60	60-80	80-100	100-120
No. ofstudents :	250	120	100	70	50

4. Given sin 45^0 = 0.7071, sin 50^0 = 0.7660, sin 55^0 = 0.8192, sin 60^0 = 0.8660, find sin 52^0 using Newton's forward interpolating formula.

5. Given $\log_{10} 654 = 2.8156$, $\log_{10} 658 = 2.8182$, $\log_{10} 659 = 2.8189$, $\log_{10} 661 = 2.8202$, find using Lagrange's formula the value of $\log_{10} 656$.

6. Fit a Lagrangian interpolating polynomial y = f(x) and find f(5)

x :	1	3	4	6
y :	-3	0	30	132

7. Find y(12) using Newton' forward interpolation formula given

x :	10	20	30	40	50
y :	46	66	81	93	101

8. Obtain the root of f(x) = 0 by Lagrange's inverse interpolation given that f(30) = -30, f(34) = -13, f(38) = 3, f(42) = 18.

9. The following data are taken from the steam table:

Temp^0 c :	140	150	160	170	180
Pressure :	3.685	4.854	6.502	8.076	10.225

Find the pressure at temperature t = 142^0 and at t = 175^0

10. From the following table of half yearly premium for policies maturing at different ages, estimate the premium for policies maturing at the age of 46.

Age x :	45	50	55	60	65
Premium y :	114.84	96.16	83.32	74.48	68.48

11. Using Newton's backward difference formula construct an interpolating polynomial of degree three and hence find f(-1/3) given f(-0.75) = - 0.07181250, f(-0.5) = - 0.024750, f(-0.25) = 0.33493750, f(0) = 1.10100.

12. From the following data find y'(6)

X :	0	2	3	4	7	9
Y:	4	26	58	112	466	922

13. Find the value of sec (31) from the following data

$\theta(\deg ree)$:	31	32	33	34
Tan θ :	0.6008	0.6249	0.6494	0.6745

14. The following data gives the velocity of a particle for 20 seconds at an interval of five seconds. Find initial acceleration using the data given below

Time(secs) :	0	5	10	15	20
Velocity(m/sec):	0	3	14	69	228