1.0 Objectives of the unit

After studying this unit, you should be able to:

- Broaden your perception of the field of analytical chemistry,
- Distinguish between various analytical techniques,
- Classify the chemical, electrical, optical, nuclear, and thermal methods of analysis,
- Classify the separation methods,
- Understand the utility of analytical techniques, and
- Highlight the development of methods of chemical analysis.

1.1 Introduction

Analytical chemistry is a sub-discipline of chemistry with interdisciplinary character. This is a first Unit of this programme. In this unit we will first learn about the analytical chemistry. We know the importance of chemicals in our life. Chemicals are the foods we eat, clothes are made up of chemicals and medicines are full of chemicals and so on. We all know the importance of chemical analyses in all these products. The development of methods for chemical analysis lies in the province of analytical chemistry. Therefore, in this unit first we will learn about the analytical chemistry and its important. Then we will study about different chemical methods of analysis like gravimetry and volumetry analyses. We will also discuss briefly about the classification of electrical methods of analysis which include potentiometry, amperometry, voltammetry, coulometry, etc. In the latter part of the unit we have introduced you to optical method of analysis like emission, absorption, UV, IR, and Raman spectroscopic. We have also discussed nuclear and thermal methods of analysis briefly. Separation of the chemicals compounds is a very important part of analytical chemistry. So we introduce you about the separation methods. There is a separate course on separation methods i.e. (course MCH-002) in this PGDAC programme, where you will study all these techniques in details. In the last you will study emerging needs and recent trends of analysis.

1.2 Analytical chemistry-an introduction

In our daily lives we cannot help but make use of chemicals. Chemicals are the food we eat, these are the clothes we wear, the medicines we take and the immense variety of articles that we use. As a chemist whenever you see a material, two curiosity questions are raised in your

mind: What is it? And how much? The answer of these questions is given by chemical analysis. In chemical analysis you are able to analyse compounds to see what and how much of a given chemical constituent they contain. However, before an analysis can be made, a method must be available. *The development of methods for chemical analysis lies in the province of Analytical Chemistry*.

In developing methods of analysis the analytical chemists tried to use the principles from any field of science: chemistry, physics, biology, biochemistry, geology, engineering and computers, etc. The constituents to be analysed may be elements, ions, radicals, functional groups or compounds. Analytical Chemistry is thus concerned with the theory and practice of methods used to determine the composition of substances. At the present time no material is taken into production or released without analytical data which characterize its quality and suitability for various purposes. The qualities of food materials and medicines, etc. all need analysis results prior to be used. These results not only form the basis of all processing calculations but they also determine the costs of materials, which form the basis of financial estimates. This makes an analytical laboratory as an important section of chemical industry.

The last 3-4 decades witnessed an exciting phase in the development of analytical chemistry, in terms of new and improved analytical techniques which permit separation, detection, structure elucidation and quantification of much lower levels of chemical species, multi-component analysis and a much short duration required for analysis. The analytical methods can be divided into three major groups. They are chemical, physical and physicochemical methods. The chemical methods are the *classical methods* whereas the physical and physicochemical methods are known as *instrumental methods* of analysis.

A method to be called as an analytical technique should be based on the measurement of a property which is related to either the nature or the amount of the substance under examination. The analytical chemist must understand the relation between the property to be measured and the nature and the quantity of the desired chemical species to be studied. The property which depends on the nature of the substance is helpful in qualitative analysis, whereas the property which depends on the amount of the substance is useful in quantitative analysis. Mostly a quantitative analysis is preceded by the qualitative analysis, while a qualitative analysis may give a rough idea of the relative amount of the constituents present in the sample. The sample is the part of the material which is examined. In analytical chemistry

we take small samples to perform the chemical analysis and these results are taken as a measure of the results of whole material.

A substance can be determined by a variety of methods, and the analytical chemist then has to select the most advantageous method available in the laboratory. For choosing a suitable technique, the analyst must know what type of samples is to be analysed, what information is needed and for what purpose. Further, the way you perform an analysis will depend on your experience, the equipment available and the preparation of the sample for analysis, the time and the cost involved. Besides this, you have to choose an analytical method on the basis of accuracy, sensitivity and selectivity. While choosing an instrumental technique you should be aware that most instrumental methods are relative methods. Therefore, they must be calibrated with standards. Usually an analytical calibration curve of instrument response versus concentration or amount of the substance is prepared prior to analysis of unknowns.

1.3 Classification of different analytical techniques

Analytical chemistry deals with the development of methods for chemical analysis. These methods depend on the measurement of some physical property. Various physical properties which are characteristics of a particular substance or its constituent can be made the basis of an analytical technique. A listing of the types of analytical techniques demands some kind of classification. Classification schemes will vary with the viewpoint the analytical chemist considers. In the broader sense the analytical techniques can be classified on the basis of type of properties in the following way.

i) Chemical methods of analysis
ii) Electrical methods of analysis
iii) Optical methods of analysis
iv) Nuclear radiation methods of analysis
v) Thermal methods of analysis
vi) Separation methods

These methods can further be classified into different techniques depending on the measurement of a characteristic property based on either the nature or the amount of the desired constituent of the sample.

1.4 Classification of Chemical Methods of Analysis

These methods are based on the primary role of a chemical reaction. In these methods the direct

measurement of mass is carried out by one of the two procedures, i.e. by weighing or by measuring volume. The corresponding methods are known as *gravimetry* and *volumetry*. Both of these being developed at early stage are now also known as *classical methods of analysis*.

1.5 Gravimetry

Gravimetric analysis means analysis by weight and pertains to all determinations wherein the final results are obtained by means of the analytical balance. Gravimetry is an accurate macro-analysis procedure which mainly depends upon precipitation of an ionic or molecular substance on the basis of a chemical reaction. The precipitate is then separated, dried and weighed in the suitable form. The amount of the desired constituent (*analyte*) is then determined by simple calculation. At undergraduate level you have estimated barium as barium sulphate following gravimetric analysis.

1.6 Volumetry

The amount of the analyte can also be found in another way by measurement of the *volume*. The method based on accurate measurement of volume of a reagent solution of accurately known concentration, taken for a reaction is known as *volumetric analysis*. The measurement of volume makes considerable saving of time. The greater speed of volumetric analysis is an important advantage of this method over gravimetry.

The volumetric analysis is characterized by a titration, hence the method is also known as *titrimetry*. In this method a known volume of the reactant substance is taken in a beaker and the titrant is added from a burette to it till the reactant completely reacts with the titrant. We call it as the equivalence point which is indicated by an indicator. The volume of the titrant required to reach the equivalence point is noted on the burette and the calculations are made to get the amount of the constituent of the reactant. The volumetric method is much simpler than the gravimetric method. However, the reaction for volumetry should be rapid, so that, practically speaking, zero time will be needed after each volume addition of titrant for the reaction to reach equilibrium.

1.7 Classification of Electrical Methods of Analysis

An electrical method of analysis also known as electroanalytical method can be defined as one, in which an electrochemical property of a solution is measured. A classification of electroanalytical methods can be made by measuring different electrical quantities, such as,

potential, current, quantity of current, resistance and dielectric constant. These methods have different names on the basis of the measurement of these quantities and are stated below.

i) Potentiometry

ii) Amperometry

iii) Voltammetry

iv) Coulometry

v) Conductometry and High Frequency Methods

1.9 Potentiometry

Analytical methods based on the measurement of potential difference across an electrochemical cell to interpret the results of an analysis are designated by the term *potentiometry*. The term potentiometry is derived from the word "potential" which is half cell potential (electrode potential) obtained by measuring voltage across an electrochemical cell with the help of a standard electrode. The result of the analysis can be computed directly from the voltage of the cell, or the equivalence point of a titration known as *potentiometric titration*. In potentiometric titrations we discuss redox titration curves based on half-cell potentials and describe the necessary procedures to obtain the sample analyte in the correct oxidation state for titrations

A special class of potentiometry where the potential of an indicator electrode is measured as a function of hydrogen ion concentration is designated as *pH-metry*. By suitably modifying the common voltmeter to high impedance mV meter and usually making use of a glass electrode as a hydrogen ion indicator electrode suitable pH- meters can be designed to measure pH instantaneously.

i) Amperometry

Amperometry designates methods involving current measurements. The term amperometry is derived from the word "ampere" which is the unit of current. Amperometric methods are generally applied to the detection of equivalence point of titration and method is known as amperometric titration and usually involves one microelectrode as an indicator electrode (which is also polarized electrode the other electrode is the reference electrode and is non-polarized). Here the current at a fixed potential is measured as a function of titrant volume. On plotting the data two straightlines with different slopes are obtained on both the sides of equivalence point.

A modification of the amperometric titration involves the use of two polarized microelectrodes and is known as *biamperometry* or the dead stop end point titration. Here the two identical microelectrodes are immersed in a well stirred solution of the sample. A small potential is applied between these electrodes and the current is measured as a function of the volume of the titrant added.

ii) Voltammetry

In voltammetry an electroactive species is consumed (oxidized or reduced) only at the surface layer of the indicator electrode in an electrolytic cell. The resulting current, due to electron transfer process, is measured as a function of applied potential. The current versus potential curves are plotted. These curves are known as the current voltage curves or current-potential curves or I-E curves. The shape of the I-E curves depends on the polarization of the indicator electrode, whereas the other electrode known as the reference electrode remains non-polarized. In voltammetry we study the relationship between the current and electrode potential and its application to chemical analysis.

When a special type of micro electrode, the *dropping mercury electrode* is used as the indicator electrode the technique is known as *polarography*. A continuous changing potential is applied across the dropping mercury electrode and the reference electrode and the resulting current is monitored by a current measuring device. The current- potential curves are known as *polarograms* or *polarographic waves*. The half wave potential ($E_{1/2}$) of a polarographic wave can be a useful quantity for qualitative identification whereas the height of the wave is used for quantitative estimation of the particular species. A recent development in polarography has been made by the use of three electrodes. This modification makes it possible to obtain sharply defined polarographic waves from non-aqueous solvents that have low electrical conductivities.

iii) Coulometry

Analytical methods based on the measurement of the quantity of electricity are designated by the term *Coulometry*. The term is derived from "coulomb", which is one of the units used for quantity of current. A fundamental requirement of all coulometric methods is that the species determined interacts with 100% current efficiency.

Coulometry is performed by two general techniques: (i) coulometry at constant potential, known as

potentiostatic coulometry and (ii) coulometry at constant current known as amperostatic coulometry.

In the first the potential at the working electrode (the electrode at which the analytical reaction occurs) is maintained at a controlled level. Here current is initially high and decreases exponentially with time. In the second, that is, coulometry at constant current, a constant current is operated for a time until a signal indicates the completion of the analytical reaction. These methods are, frequently called coulometric titrations. The second type of methods has enjoyed wider applications than the first.

iv) Conductometry and High Frequency Methods

The measurement of conductance (the reciprocal of the resistance) can sometimes be useful in chemical analysis. Methods based on electrical conductance measurements are grouped under the term *conductometry*. Conductometry is performed in two different ways: in one an analysis can be computed directly from conductance measurements and in other conductometry is applied to the determination of the equivalence point of titrations (conductometric titrations). Conductometric titrations where the change in conductance is related to concentration changes of the ionic species involved in the titration reaction are more frequently applied. In conductometry conductances are generally measured by using alternating current of3-6 volts with frequency of 50-1000 Hz. The technique has the advantage of simplicity and good sensitivity.

In conventional conductometric method (as stated above) conductance measurements are made by using alternating current at relatively low frequencies (50 – 1000 Hz).

The technique can be modified by using much higher frequencies (several mega Hz) and the methods are called as *high frequency methods*. These methods can be applied to the measurement of dielectric constant and also for several titrations where the electrodes do not come in the intimate contact of the solution.

1.10 Classification of Optical Methods of Analysis

You have studied the classification of electroanalytical techniques. In this subsection you will study the classification of optical methods of analyses briefly. In course MCH-003 of this programme you will study Optical Methods of Analysis in details. These methods are now called as spectroscopic methods of analysis. In these methods the first instruments were

developed for use of visible region and therefore called optical methods. This term has by now been extended to other regions of emr as well. All spectroscopic methods are based on the interaction of electromagnetic radiation with the quantized energy states of the matter. Here we study the measurement of a quantity based on emission, absorption, scattering or change in some property of electromagnetic radiation (emr) depending on the nature or the amount of the constituent of the sample. The classification may be based on either the type of effect (emission, absorption or scattering) or the type of the emr (x-ray, uv- vis, IR etc.) used. The important spectroscopic methods are mentioned below.

i) Emission Spectroscopy
ii) Absorption Spectroscopy
iii) Ultraviolet and Visible Absorption Spectroscopy
iv) Infrared Absorption Spectroscopy
v) Fluorophotometry
vi) Turbidimetry and Nephelometry
vii) Raman Spectroscopy

i) Emission Spectroscopy

Emission spectroscopy is the method where the characteristic spectrum produced by excitation of elements is applied to qualitative and quantitative analysis. These methods depend on the electromagnetic radiation produce when the analyte is excited by thermal, electrical or radiant energy. Each element has a characteristic emission spetrum, this is applied to qualitative analysis. Quantitative determinations are also possible, as during the burning of the sample under controlled conditions the energy emitted for a given spectral line of an element is proportional to the number of atoms that are excited and consequently to the concentration of element in the sample.

ii) Absorption Spectrometry

Absorption spectrometry is based on the measurement of the absorption of electromagnetic radiation by matter. Absorption refers to a process by which a chemical species in a transparent medium selectively absorbs the photons of certain electromagnetic radiation. The absorption varies with the wavelength of incident radiation. Measurement of absorption can be made at

single wavelength or over a wide range of wavelengths. When the selection of a narrow wavelength range is made with the help of filters the analytical technique is known as *filter photometry*. When the measurements are made for approximately monochromatic radiation obtained with the help of a monochromator (such as: prism or grating) the technique is known as *spectrophotometry* or *absorption spectrometry*.

Absorbances are easily measured in each spectral region and are of great utility in analytical studies. The various types of spectroscopic methods are also named on the basis of the spectral region used.

iii) Ultraviolet and Visible Absorption Spectroscopy

Analytical methods which involve the measurement of absorption of *ultraviolet and visible radiation (wavelength range from 180 to 780 nm) by an atomic, ionic, or molecular species are known as ultra violet and visible spectroscopic methods*. UV and Visible spectroscopy involves transitions between electronic levels of absorbing chemical species. These methods find application in qualitative as well as quantitative analysis. For qualitative analysis uv-visible spectra provide a valuable tool in the identification of unsaturated organic compounds and elucidation of their structure.

However, these methods particularly in the visible range are mainly used for quantitative analysis of substances of different categories.

iv) Infrared Absorption Spectroscopy

Infrared absorption spectroscopy involves the absorption of infrared radiation (wavelength range from 0.78 to 1000 □m) depending on increasing the energy of vibration or rotation associated with a covalent bond, provided that such an increase results in a change in the dipole moment of the molecule. Infrared (IR) spectroscopy can further be subdivided into near IR, middle IR and far IR spectroscopy. Majority of applications lies in the middle IR region. IR spectroscopy finds widespread application to qualitative and quantitative analyses. However, its most important use has been for the functional group identification of organic compounds.

v) Fluorophotometry

The energy of the photons of incident radiation absorbed and changes the absorbing species to excited state. Certain chemical substances (known as photluminiscent) after excitation can re-emit radiation. Re-emission of radiation can be immediately ($< 10^{-8}$ sec) after the absorption and is known

as *fluorescence*. The fluorescence intensity is practically proportional to the concentration of fluorescent substance. The measurement of the intensity of fluorescence serves useful analytical purposes and thetechnique is known as *fluorophotometry*.

When re-emission of radiation takes longer time (minutes, hours, or days) thephenomenon is known as phosphorescence and the related technique as *phosphorimetry*.

vi) Turbidimetry and Nephelometry

These methods are applied to determine the concentrations of suspensions where smallsolid particles are homogeneously dispersed in the liquid medium. Analytical method where determinations are made by measuring opacity of suspension of small particles with the help of measuring intensity of transmitted light is known as *turbidimetry*.

The analytical method which is based upon the measurement of intensity of light scattered by a suspension of small particles is designated as *nephelometry*.

Turbidimetric and nephelometric methods frequently give erratic results, therefore these methods are applied only when the results need not to be very accurate and other precise methods are not available.

vii) Raman Spectroscopy

Raman spectroscopy involves the scattering of electromagnetic radiation by a liquid (solution) following **Raman effect** (scattering with change of wavelength). The shiftin wavelength in Raman effect is caused due to extraction of energy from the quanta of incident radiation and utilize to raise molecules to higher vibrational states. The scattered radiation thus has less energy and higher wavelength. Raman and infrared techniques concern vibrational energy change and they are complimentary to each other. An important advantage of Raman spectra over IR spectra lies in the fact that water does not interfere in Raman spectroscopy and aqueous solutions can be handledvery well. Some other optical methods, namely, flame photometry, refractometry, polarimetry find applications in analytical laboratory but are not considered due to their less importance.

1.11 Classification of Nuclear Methods

We shall now examine some techniques which can provide analytical information based on nuclear properties. Each of these properties or combinations of them can be studied suitably by analytical chemistry. Nuclear method can be group into following.

i) Radiochemical Methods

ii) Radiometric Methods

iii) Isotopic Dilution Methods

iv) Activation Analysis

v) Mossbauer Spectroscopy

vi) Nuclear Magnetic Resonance Spectroscopy

vii) Mass Spectrometry

Table 1.1 lists these methods along with the property measured.

Table 1.1: Nuclear Methods of Analysis

S.N	Name of the method	Property measured	Mechanism involved
1.	Radiochemical methods	Radioactivity	Radioactive disintegration of radioisotopes can be measured with high sensitivity and specificity.
2.	Mossbauer spectroscopy	Rasonance absorption of γ-rays	Resonance fluorescence of γ-rays and involves intranuclear energy levels.
3.	Nuclear magnetic rasonance spectroscopy	Position of signals (chemical shift) and their intensity in NMR spectrum	Interaction of quantized nuclear spin with an applied magnetic field
4.	Mass spectrometry	Position and intensity of signals of mass spectrum	Mass to charge ratio of ionized atoms or molecules

1.12 Radiochemical Methods: Very small amounts of radioactive substances (having natural or artificial radioactive isotopes) have measurable activity, this makes the development of sensitive analytical methods known as radiochemical methods. Radiochemical methods are classified into

three categories, namely radiometric analysis, isotopic dilution methods and activation analysis. All thesemethods have high sensitivity, specificity and good accuracy.

Radiometric Methods: In *radiometric methods* a radioactive reagent is employed to separate the analyte completely from the bulk of the sample, the activity of the isolated material is then measured. In an alternate way when a radioactive reagent is used to litrate the analyte and the end point is established by activity measurements, the method is known as radiometric titration.

i) **Isotopic Dilution Methods**: In an isotopic dilution method a known amount of the same substance containing an active isotope is added to the unknown and thoroughly mixed with it. A sample of the pure substance is then isolated from the mixture and its activity is measured. The quantity of the substance in original material is then determined by simple calculation. The technique is suitable where a compound can be isolated in pure state, may be, with only a poor yield. Isotopic dilution method has the advantage of not requiring quantitative separation of the analyte.

ii) **Activation Analysis**: The methods of *activation analysis* are based upon the measurement of radioactivity induced in the samples by irradiation with suitable particles such as neutrons, protons, deutrons or helium – 3 ions. However, thermal neutrons from a nuclear reactor are the most commonly used particles and this technique is known as neutron activation analysis.

iii) **Mossbauer Spectroscopy**: The analytical methods based on the study of thephenomenon of the resonance fluorescence of gamma rays are known as *Mossbauer spectroscopy*. It involves intranuclear energy levels. An important characteristic of this radiation is the extreme sharpness of lines.

iv) **Nuclear Magnetic Resonance Spectroscopy**: *Nuclear magnetic resonance* (NMR) spectroscopy is based upon the measurement of absorption of electromagnetic radiation in radiofrequency region by nuclei of atoms of certain elements (isotopes) in the influence of strong magnetic fields. Since a nucleus bears a charge its spin gives rise to a magnetic field, the resulting magnetic dipole is oriented along the axis of spin and has value based on the nature of the nucleus. As a consequence certain atomic nuclei on exposure to a magnetic field would lead to splitting of their energy levels. The position of the signals of NMR spectra is characterized in terms of chemical shift which is very useful instructure elucidation.

An analogous method called *electron spin resonance* (ESR) is based upon the absorption of

microwave radiation by an *unpaired electron* when it is exposed to a magnetic field.

v) **Mass Spectrometry**: *Mass spectrometry* is a method based on the study of converting molecules into charged particles (molecular ions), their separation on the basis of mass to charge ratio and measurement of the relative intensity of lines of mass spectrum. The instrument used is a mass spectrometer which separates the charged gas molecules (ions) according to their masses.

Mass spectrometry is capable of providing qualitative and quantitative information about either the atomic or the molecular composition of the sample. It is an important tool for elucidating the structure of organic compounds.

1.13 Classification of Thermal Methods of Analysis

In thermal methods of Analysis some property of the system is measured as a function of temperature. In some of these methods the temperature is used as an independent variable while in some others as a dependent variable say time. The recorded curves are helpful in interpreting the thermal behaviour of the sample.

Thermal methods are classed into nearly a dozen varieties. Out of these some commonly used methods are:

i) Thermogravimetric Analysis (TGA)

ii) Derivative thermo-gravimetry (DTG)

iii) Differential Thermal Analysis (DTA)

iv) Differential Scanning Calorimetry (DSC)

v) Thermometric Enthalpy Titrations (TET)

Property measured and instrument used for these methods are given in Table 1.2.

Table 1.2: Thermal Methods

S.N	Name of the method	Property measured	As a function of	Instrument
i	Thermogravimetric	Change in weight	Temp	Thermobalance

	Analysis (TGA)			
ii	Derivative thermo-gravimetry (DTG)	Rate of change in weight	Temp	Thermobalance
iii	Differential Thermal Analysis (DTA)	Heat absorbed or evolved	Temp	DTA apparatus
iv	Differential Scanning Calorimetry (DSC)	Thermal Transition	Temp change	DSC cell
v	Thermometric Enthalpy Titrations (TET)	Temp. change	Volume of titrant	Titration calorimeter

1.13 Thermogravimetric Analysis (TGA)

Thermogravimetric analysis (TGA) involves the measurement of mass of a sample as its temperature is increased at a linear rate. In TGA a suitable solid sample undergoes the following type of reaction;

$$\text{Reactant(s)} \longrightarrow \text{Product(s)} + \text{Gas}\uparrow$$

causing a loss of mass due to gas being evolved. Plot of mass versus temperature known as *thermogram* persuits determination of thermal stabilities and sample compositions at different temperatures.

Derivtive Thermogravimetry (DTGA)

Derivative thermogravimetric analysis (DTGA), as the name suggests, is carried out by the plot of first derivative of thermogram. In DTG a plot is prepared in dw/dT versus T. In DTG curves the changes are observed in the form of maxima or minima which make the weight changes very clear even those which are not so clear in thermograms.

Differential Thermal Analysis (DTA)

In *differential thermal analysis* (DTA) the difference in temperature between sample and reference is measured as a function of temperature. From the plot of ΔT versus T. the transition temperature and the nature of the change (exothermic or endothermic) can be determined.

i) Differential Scanning Calorimetry (DSC)

In *differential scanning calorimetry* (DSC) the sample and the reference material are subjected to precisely programmed temperature change. When a thermal transition occurs in the sample, the temperature changed in balanced by adding thermal energy to either the sample or the reference. DSC can measure directly both the temperature and enthalpy of a transition.

ii) Thermometric Enthalpy Titrations (TET)

The *thermochemical methods* where a titration process is followed by the measurement of enthalpy change during the course of a reaction carried outunder controlled conditions usually in a small Dewar flask, are known as *thermometric enthalpy titrations*.

The above lists do not conclude all the methods of chemical, physical and physicochemical methods of analyses. Useful analytical determinations can be made on the measurement of various other properties such as refraction, reflection, surface tension, polarization of light, etc. However, such methods are not listed here just for the sake of convenience.

1.14 Classification of Separation Methods

In the previous subsections 1.3.1 – 1.3.5 you have learnt that the determination of a substance which is free from interfering substances can be accurately made by the direct application of the suitable technique. However, in natural samples the state of affairs is likely to be much different because of their complex nature by the presence of interfering substances. Extremely few methods may be specific or even selective and accuracy in determinations by most methods are affected by the interfering substances. It is, therefore, frequently necessary to perform quantitative separations with the objective either for isolation of the analyte or to remove the interfering substances. Therefore separation is a prerequisite procedure for such determinations. Though separation is not a purely analytical technique but it is commonly required prior to many analyses.

You will know about some methods of separations in this subsection. In separations, in general by appropriate reactions, the desired constituent is brought into one phase

and interfering elements are brought into another and the phases being separated by physical processes. Some methods of separation are the following:

A. Classical methods

a) Precipitation

b) Distillation

c) Sublimation

d) Formation of complexes

B. Modern methods

i) Chromatography

ii) Solvent extraction

iii) Ion-Exchange

iv) Electrophoresis

You are well aware about the classical methods and only the modern methods will be defined here briefly and in course MCH-002 you will learn these methods in details.

Chromatography

Choromatography is a multistage separation process in which the sample is applied on a stationary phase over which a mobile phase is percolated. Various solutes present in the sample are separated on the basis of differential migration. Chromatography can be classified into various kinds depending on the nature of stationary and mobile phases and the mechanism of distribution involved. These kinds are named as paper chromatography, thin-layer chromatography, liquid chromatography, high performance liquid chromatography, gas chromatography, gel chromatography, partition chromatography, adsorption chromatography, ion exchange chromatography, electrochromatography etc.

Chromatography has been used with remarkable success in the separations of inorganic, organic and biochemical substances. The separations of vitamins, hormones, natural pigments, fission products of uranium and steroids etc.are

some good examples of its scope and success.

Solvent Extraction

In *solvent extraction,* a desired solute can be isolated/extracted by distributing it between two immiscible liquids. It exploits the differential solubility of a given solute in two immiscible solvents to separate it from the given mixture. Solvent extraction can be applied as a single stage procedure or a multistage procedure (counter current extraction).

Ion Exchange

Ion exchange is a stoichiometric process in which a solid (insoluble) material, known as ion exchanger, when comes in contact with an electrolyte solution takes either positive or negative ions (known as counter ions) and releases the ions of like charge (to maintain the stoichiometry) to the solution. The solid materials having cations as exchangeable ions are known as *cation exchangers* and having anions as exchangeable ions are known as *anion exchangers*. Ion exchange is a reversible process. The exchanged ions can be replaced by other ions of like charge. Ion exchangers find great utility in separating the ionic species of similar nature. Some separations of common interest are of rare earthelements and of amino acids.

Electrophoresis

The movement of charged particles in the influence of an electric field, in general, is known as *electrophoresis*. If the components of a mixture have different velocities under the influence of the electric field, it is possible to separate them. This method has been used with remarkable success for the separation and characterization of polysaccharides, nucleic acids, haemoglobins and other high molecular weight compounds. Small organic and inorganic ions can also be separated with the help of this technique (known as *ionophoresis*).

SAQ 1

What is the essential feature of a method to be called as an analytical technique?

……...

……...

……...

……...

……...

SAQ 2

Name the methods that are now known as classical methods.

……...

……...

……...

SAQ 3

Define polarography.

……...

……...

……...

……...

SAQ 4

Is chromatography a single stage or a multistage separation process?

……...

……...

……...

1.15 Criteria for evaluating the utility ofanalytical techniques

Analytical chemistry is of enormous importance in science and industry. It deals with the development of methods for chemical analysis which are utilized in detection, determination and separation of chemical constituents and structure elucidation of chemical compounds. For example, the chemical formula of an unknown substance is ascertained from the percentage contents of its constituents found by analysis. Today,

with the help of newer techniques, such as mass spectrometry, NMR spectroscopy, high performance liquid chromatography, etc., the structure elucidation has become more perfect.

Utility of analytical techniques is to be found in various fields as the key to the solution of a variety of scientific problems and of industrial problems. It is a chemical discipline with interdisciplinary character providing valuable information in many branches of science and technology. All fundamental laws are based on analysis.

Mechanisms in so many chemical reactions are developed as a result of analysis. Much of what is known of the mechanisms by which chemical reactions occur has been learned through kinetic studies employing quantitative measurements of the rates at which reactants are consumed or products are formed. You understand the importance of rates in an industrial process to decide the cost of the products coming out of an industry.

The results of a typical quantitative analysis are based upon two series of measurements, one of which is related to the amount of sample taken, and the second to the relative amount of the desired constituent present in the sample. On the basis of the amount of the sample taken the methods are named as macro, meso, micro, and ultramicro methods. On the basis of the relative amount of the desired constituent the results take the form of numerical data in suitable units such as percent, parts permillion, parts per billion or some other form.

In order to understand the criteria for evaluating the utility of the analytical techniques, it is useful to identify the several steps in performing quantitative analyses. A complete analysis actually consists of the following main steps:

i) Sampling

ii) Dissolution of the sample

iii) Separation of interfering substances

iv) Measurement

v) Interpretation of the measurements

Often beginners carry out usually steps iii, iv and v, for the sake of convenience.

Sampling

The heart of the quantitative analysis is to carry a sample with great care through a number of manipulations without accidental losses and without introducing foreign material, since the sample is representative of all components and their amounts as contained in the bulk material. Conclusions will be drawn about the composition of the bulk material from the analysis of a very small portion of the material.

Aknowledge of statistics is of considerable importance as an aid to establishing sampling programmes so that data obtained may be subjected to statistical treatment when necessary.

Sampling techniques may be quite different in different cases. Each type of material has its own special sampling instructions which take into account the specific characteristics of the material, the quantity taken, its purpose, etc.

Dissolution of the Sample

Most analyses are performed on solutions of the sample. Therefore, suitable solvent is required to dissolve the sample rapidly and under conditions in which there is no loss of the analyte. The dissolution process depends on the nature of the sample material.

Two most common methods employed in dissolving inorganic sample are (1) treatment with hydrochloric acid, nitric acid, mixture of hydrochloric and nitric acids, sulphuric acid or perchloric acid, and (2) fusion with an acidic or basic flux followed by treatment with water or an acid.

Organic solvents are preferentially taken to dissolve the samples of organic nature. However, special methods are to be developed to dissolve a silicate material, a high molecular weight polymer or a specimen of animal tissue.

Separation of Interfering Substances

The interfering substances are the compounds or elements that prevent the direct measurement of the species being determined. Therefore, before an analytical measurement can be made it is usually necessary to solve the problem of

interferencesby their separation from the analyte.

There may be two ways in general to achieve separation (i) by isolating the desired constituent in a measurable form or (ii) by removing the interfering substances fromthe desired constituent.

In most of the separation techniques the substance of interest is transferred from one phase to another. Therefore the separation procedures can be classified depending onthe type of phases involved in these procedures. There may be four suchcombinations: solid-liquid, liquid-liquid, solid-gas and liquid-gas.

1.16 Measurement

The way of measurement depends upon the type of analytical technique being used. A gravimetric method involves the measurement of weight of a suitable form of the analyte. In a volumetric method the measurement is of the volume of a solution of known concentration which is required to react with the analyte. A characteristic feature of most instrumental methods is the necessity for finding empirically the value of the intensity factor corresponding to the mass or concentration of a given constituent. The majority of these methods therefore, require calibration by the use of a standard containing a known amount of constituent, which serves as a basis for comparison in the measurement.

The established procedure must be followed carefully in individual quantitative determinations. Equal care must be taken in the choice of suitable technique for the desired determination. The success or failure of an analysis is often critically dependent upon the proper selection of method.

The analytical chemistry can save much time and improve the accuracy of results by a critical comparison of the various methods on the basis of certain criteria. Although no simple rule can be prescribed but some essential characteristics of a method which must always be considered are the following:

a) The complexity of the materials to be analyzed

b) The probable concentration of the species of interest

c) Accuracy

d) Sensitivity and detection limit

e) Selectivity

f) Duration of an analysis

g) Cost of equipment

Interpretation of the Measurement

The result of an analytical measurement after proper calculation is usually reported inrelative terms, that is, in some way that expresses the quantity of the analyte present per unit weight or volume of the sample. Thus the results take the form of numerical data in suitable units such as percent, parts per million or some other.

The methods of statistics are commonly used and are especially useful in expressing the analytical results. You should remember that the analytical results can be reliable only if all the conditions for which the particular method was developed and verified are strictly obeyed. Any deviation from these conditions leads to error and loss in accuracy.

SAQ 5

Name any two criteria helpful in comparing the analytical methods.

……...

……...

……...

……...

1.17 Emerging needs and recent trends

Analytical chemistry is of fundamental importance as it is based on the development of methods for identification, determination and separation of substances in different ranges – be it a major or a trace constituent. Identification of a new compound and its structure elucidation, for example, need the result of chemical analysis for its conformation. It is impossible to understand a chemical reaction without knowing the quantities involved. The study of the mechanisms of all chemical reactions at one stage or the other requires the use of analytical chemistry. Side by side with its use in the

development of fundamental aspects of science analytical chemistry finds its needin the day to day life of the people. Today there is a great need of chemical analysis in industry, since it provides the means of testing raw materials and for assuring the quality of finished products. Most industrial products such as fuels, paints, pharmaceuticals, food materials etc., cannot be sold to the consumer without the analysis information written on the label of the product. It is the analytical chemistry which decides, whether or not, an ore can be used for profitable extraction of an element, whether a product manufactured in an industry is under quality control or not. Not only in chemistry but in such diverse branches of science as physics, biology, mineralogy, medical sciences, geology or technology etc, analytical chemistry is required in the solution of many intucate problems.

Analytical chemistry is recognized to fulfil both the fundamental as well as the applied aspects of science and technology in general and chemistry in particular. The explosion of industrial and technological developments in recent years has created analytical problems which demand increasingly sophisticated knowledge and instrumentation for their solution. You might have heard about a number of problems where the determination is required at parts per million or even parts per billion level. Some typical examples of such problems can be mentioned as: determination of impurities at parts per billion levels in semiconductor materials, determination of pesticide residues at the trace level in food products, detecting traces of pollutants in the environment, analysing a blood sample, or analysing a giant protein molecule for several amino acids present in it. The solutions to such problems have successfully been developed by analytical chemists.

At early stage you have gone through the study of the fundamental laws of chemistry. You know that they are based on quantitative analysis. The earliest analytical determinations systematically developed were based on the determination of the massof a constituent. *Antoine Lavaisier* has been considered the "*father of analytical chemistry*" because of the careful quantitative analysis he performed on conservationof mass (using the analytical balance).

Historically, the first quantitative analyses were gravimetric made possible by the

invention of precise balance. Gravimetric analysis means analysis by weight, was properly developed in the seventeenth century.

Soon after the development of gravimetric analysis, it was found that carefully calibrated glassware made possible considerable saving of time through the volumetric measurement of standardized solutions. This technique is known as volumetric analysis or titrimetry.

Both gravimetry and volumetry are capable of high accuracy but are valuable for major constituents. These methods are inherently simple, involve no prior calibration, no investment in expensive equipment and no high degree of specialized knowledge. These methods remained for many years the only quantitative procedures available for nearly all analyses. These methods are now known as classical methods of analysis.

The methods developed mainly in the last century are known as modern methods of chemical analysis. In general, a technique at early stage is first developed, a few people become familiar with the new technique, they develop it for their own ends and the new methodology spreads throughout their field. The modern methods are classified into two groups; (a) non-instrumental, (b) instrumental. The non-instrumental methods are mainly the methods of separation that are required prior to analysis and the important ones are: chromatography, ion exchange, electrohoresis and solvent extraction. Although there is no clear demarcation between non-instrumental and instrumental methods, the difference is just of degree. The instrumental methods are more sophisticated as the instruments are composed of rather complicated electronic, optical and mechanical parts. In instrumental techniques the principles of physics and physical chemistry are applied to study a particular property as a function of nature or amount of the substance of interest. The function of an instrument is to translate the chemical composition into information directly observable by the operator. The interest in instrumental analysis has been increased, mainly, because of two reasons (a) they greatly reduce the duration of many analyses which are quite tedious and much time consuming by classical methods, (b) they can be applied to the determination of substances under conditions in which classical methods fail (e.g. trace analysis).

Instrumental methods are developed only recently mainly because of two reasons: firstly

the physical properties on which these methods are based were not known earlier, secondly the design of the instrument required a lot of investment.

Furthermore, for every physical property to be measured, a separate instrument is to be designed, e.g., a spectrophotometer to measure absorbance, a pH-meter to measure pH and a conductometer to measure conductance. The invention of spectroscope brought with it an extremely fruitful analytical approach and gradually a few colorimetric methods were investigated. After that it was found that electrical measurements could detect end points in titrations and several electroanalytical methods were developed. Somewhere in the middle of the last century the rapid development of sophisticated instrumental components, vacuum-tube amplifier, photoelectric tube, phtomultiplier tube, transistors, semiconductor devices etc. has resulted in the establishment of many analytical instruments based upon them.

The recent trends in analytical techniques have followed closely the development of new measuring instruments to find faster and convenient ways for sensitive and selective determination of desired constituent. Today we have a number of instrumental methods which can analyze in minutes what was done in days before. Not only the time has been shortened but with the help of new techniques it is now possible to detect and determine trace constituents at micro and nano or still lower levels, and to analyze mixtures which could not be analyzed by the earlier methods.

Thus, there have been three major contributions in the present vitality of analytical chemistry

i) The flow of theory from physical chemistry and physics into analytical chemistry,

ii) Application of electronics by analytical chemists for assembling new, faster and more sensitive instruments,

iii) Application of computers for both data processing and automated control of instruments.

Such contributions have resulted in an explosive development in regard to the kinds of samples to be analysed and the sensitivity of the identification and determination of the analyte. This also resulted in a great decrease in time and human labour

required for each determination.

One of the recent developments in analytical chemistry during the last few decades has been achieved by the appearance of commercial automatic analytical systems. Here the ready made analytical data is provided in a much smaller time with minimum efforts of the operator. Initially these systems were applied for routine analyses of clinical laboratories. Now, they find their utility in diverse fields of routine chemical analysis and control of industrial processes.

Often the analysts have to handle a large number of samples and/or require to process vast amounts of data. Instruments can be designed that will automatically perform many or all the steps of an analysis. Computer techniques can be applied or even be interfaced to the analytical instruments and the results can be automatically interpreted as desired. Such automation is very helpful to get continuous online analytical information of the quality control of an industrial plant process.

Now it is possible to perform all steps of an analysis starting from sampling through measurement to data display using automatic analytical instruments of new sophisticated design. These devices make the analysis very rapid either as continuousprocess or discrete analysis.

SAQ 6

i) Who is considered as the father of Analytical Chemistry?

ii) When gravimetric analysis was developed?

……...

……...

……...

……...

1.18 Summary

In this unit you learnt about the general perspective of analytical chemistry. Analytical chemistry is concerned with the detection, determination and separation of substances. It deals with the development of methods for chemical analysis.

Applications of chemical analyses are to be found every where in a scientific and industrialized society. Methods of chemical analysis developed at early stage (gravimetry and volumetry) are now known as classical methods, while those developed mainly in the second half of the last century are known as modern methods of analysis. Analytical methods can be classified in various ways according to the criterion adopted. The major classification gives chemical, electrical, optical, nuclear, thermal and separation methods. These methods can further be classified according to the quantity being measured. Utility of analytical techniques is to be found in various fields as the key to the solution of a variety of scientific problems. It is a chemical discipline with interdisciplinary character providing valuable information in many branches of science and technology. The recent trends in analytical techniques have followed closely the development of new measuring instruments to find faster and convenient ways for sensitive and selective determination of desired constituents. Instruments can be designed that will automatically perform many or all the steps of an analysis.

1.19 Terminal questions

1. Define (a) turbidimetry, (b) nephelometry.
2. What are the major contributions in the present vitality of analytical chemistry?
3. What mechanism involved in radiochemical methods?
4. Name the commonly used Thermal methods.

Answers

Self Assessment Questions

1. A method to be called as an analytical technique should be based on the measurement of a property which is related to either the nature or the amount of the substance under examination.
2. i) Gravimetric analysis, and ii) volumetric analysis.
3. In voltammetry when a special type of microelectrode, the dropping mercury electrode, is used as the indicator electrode, the technique is known as polarography.
4. Chromatography is a multistage separation process.

5. i) Accuracy and ii) Sensitivity and detection limit.
6. i) Antoine Lavaisier and ii) 17^{th} Century

Terminal Questions

1. a) *Turbidimetry*: The analytical method where determinations are made by measuring opacity of suspension of small particles with the help of measuring intensity of transmitted light is known as Turbidimetry.

   b) *Nephelometry*: The analytical method which is based upon the measurement of intensity of light scattered by a suspension of small particles is designated as nephelometry.

2. There have been three major contributions in the present vitality of analytical chemistry.

   i) The flow of theory from physical chemistry and physics into analytical chemistry.

   ii) Application of electronics by analytical chemists for assembling new, faster and more sensitive instruments.

   iii) Application of computers for both data processing and automated control of instruments.

3. Radioactive disintegration of radioisotopes can be measured with high sensitivity and specificity.
4. i) Thermogravimetric Analysis (TGA)

   ii) Derivative thermo-gravimetry (DTG)

   iii) Differential Thermal Analysis (DTA)

   iv) Differential Scanning Calorimetry (DSC)

   v) Thermometric Enthalpy Titrations (TET)

Unit-2

Statistical analysis

2.0 Objectives of the Unit

After studying this unit you are able to

2.1 Introduction

2.2 Errors and their causes

The measured value of a property will never be the accurate value of the property. The difference between the two is called error.

Definition of Errors

An error may be defined as the difference between a measured value and the true value of a property. Mathematically, $X = (0 - T)$ where X is the error in the experiment, 0 is the observed value,

T is the true value. Errors are generally expressed relatively as

2.3 Classification of Errors

Several factors introduce error into the measured value of a property. The errors, which creep in analytical measurements, are broadly classified as:

Determinate errors or constant errors or systematic errors

Indeterminate errors or unsystematic or random errors.

2.3.1 Determinate Errors

Those errors, which can have definite values and assignable causes, are termed determinate errors. These errors can be either determined, avoided or corrected. These errors are therefore called systematic errors.

2.3.2 Clarification of determinant error

Determinante errors can be classified into constant error and proportionate error as follows.

Constant errors: Constant errors are those in which the absolute error is independent of sample size (weight). Consider, for instance, a sample of l0 mg weighs l0.5 mg due to constant misuse of uncalibrated weight. A constant error of 0.5 mg will be introduced every time when a measurement is made in this faulty weight. It is to be noted that constant errors introduce the same absolute error in a measurement but the relative errors increases as the sample size is decreased as seen from the following example.

Suppose a precipitate is shown to have a weight of 500.5 mg when its measurement is made with a faulty weight. Let the actual weight of the precipitate be 500 mg.

If we take a smaller quantity of precipitate which is shown to have a weight of 50.5 mg when measurement is made with a faulty weight. Let its actual weight be 50 mg. In this case also the absolute error = 50.5 – 50 = 0.5 mg as the same faulty weight is used.

However, the relative error

Thus we may conclude that the error become more serious as the quantity of the sample to be measured decreases.

Proportionate errors:

Proportionate errors are those errors in which the absolute error increases in direct proportion to the sample quantity (weight). One cause of such error is the presence of impurities in chemicals used in analysis or impurities present in the sample to be analyzed as explained below.

For example, in the analysis of copper, Cu^{2+} oxidizes I^{-}, so that the iodine liberated is titrated against sodium thiosulphate. If ferric ion is present as contaminant in cupric salt, it also oxidizes I^{-} to iodine so that analysis of copper will yield a high percentage of copper. If the quantity of the sample is doubled the amount of iodine liberated by Cu^{2+} and Fe^{S+} will also be doubled. Hence, the absolute will also be doubled. However, the relative error will be the same, since the absolute is proportional to the

quantity of the sample reported percentage of copper will depend on the quantity of the contaminated sample.

From the above, we may conclude that when error in the experiment are such that absolute error is the same but relative error varies, the errors introduced are constant error. On the other hand, when absolute error varies and relative error remains constant, the error introduced is proportionate error. Besides the above, the following characteristics of determinant error to be noted. They may also be variable but of such a nature that can be accounted for correction. These are generally unidirectional with respect to the true value (either high or low). Determinate errors are reproducible and can be predicted by a person who thoroughlyunderstands all the aspects of the measurement.

2.3.3 Causes of Determinate Errors

The determinate errors may arise due to various factors and in general can be classified into the following four types depending upon the system measured, observer and the instrument used.

- Instrumental errors
- Methodic errors
- Operational errors
- Personal errors or Human errors

Instrumental error

These errors arise due to faults in the tools used by the analyst. Glasswares such as pipettes, burettes, measuring flasks, etc. used in volumetric analysis, instruments using electronic circuits, such as pH meter are all calibrated at a certain temperature.

If these are used in some other temperature, the calibration is disturbed and the measurements become unreliable. So instrumental errors are due to the use of uncalibrated weights, ungraduated glassware and other instrumentssuch as pH meter. Errors may also be introduced due to voltage fluctuation in the source of current supply.

Methodic error

Adoption of defective experimental methods causes these errors. These may arise due to incompleteness of a reaction and incorrect sampling. For example, Kjeldahls method used for the determination of nitrogen may not give consistent results in certain cases as in case of some organic compounds, containing ring nitrogen. The digestion with concentrated H_2SO_4 may not completely convert the ring nitrogen to ammonium sulphate. This is particularlytrue for pyridine compounds in which the

results of nitrogen determination are low. Some sources of methodic errors include coprecipitation, postprecipitation, of impurities, side reaction, slight solubility of precipitate, impurities in reagent, etc. These are most serious errors. These are inherent in the method and cannot be minimized or corrected unless the conditions of the determination are changed.

Operational error

These errors arise due to lack of knowledge or total ignorance of handling equipment and not taking the necessary precautions in measurements as exemplified below.

Lack of experience of the analyst resulting error in weighing and volume readings.

Introduction of foreign materials in the experimental sample due to carelessness such asnot covering the sample container.

Errors may be due to defective operations such as during transfer of solution, incomplete drying of samples and bumping during sample dissolution, etc.

Weighing a crucible when it is hot and cooling in a desiccator with a poor desiccant.

The use of indicators in quantities is more than necessary. This leads to erroneous results,since the indicator may also get titrated.

Ignition of precipitate in incorrect temperature.

Allowing hygroscopic materials to absorb moisture before or during weighing.

Failure to apply buoyancy correction when required.

Frequently the sources of an error may lie in more than one of these categories. For example, some error may always be expected in weighing a hygroscopic substance, but it may be increased further if the analyst has a poor balance technique.

Personal error or human error

These errors are due to factors for which the individual analyst is responsible and are not connected with the method or procedure. These errors may arise as a consequence of faulty ideas, improper technique, carelessness, ignorance and physical limitations of the experiments. Such error may be due to physical disability like colour blindness, which may make incorrect judgement of colour. Some examples of personal errors are

Mechanical loss of material in various steps of analysis.

Errors in reading in burette.

Improper washing of a precipitate.

Insufficient cooling of crucible.

Using impure reagent.

Mathematical error in calculation.

By an appropriate choice of equipment, calibration of apparatus used, and the method of analysis, systematic error can be minimized to an acceptable level.

2.4 Indeterminate Errors

These are random or accidental errors which arise from uncertainties in a measurement that are unknown and not controlled by the analysts. These are revealed by small differences in values of successive measurements made under identical condition by the same analyst. These errors follow a random distribution. Thus the mathematical law of probability can be applied to get most probableresult from a series of experiments made. A normal distribution curve (Gaussian curve) is shownin Figure 4.l for such errors which shows that

Small errors occur more frequently than large one.

Positive and negative errors of the same magnitude are likely to occur equally.

Narrow peaked curves with steep slopes indicate a relatively high precision.

Broad curve indicates a relatively low degree of precision.

These errors can not be attributed to any known cause. They are random in nature and lead to both high and low results with equal probability. They cannot be eliminated or corrected and arethe ultimate limitation on the measurement.

Propagation of errors

The uncertainty in each measured value such as weight, volume, length etc. (measured twice or more) must be estimated. This can be done according to the following rules of propagation of errors.

Uncertainty Involving Addition and Subtraction

If the measured values are added or subtracted to get the result, the uncertainty in the results is equal to the sum of the uncertainties of the individual measured value. If the measured values are$a$ and b and their uncertainties are $\pm\Delta a$ and $\pm\Delta b$ respectively, then

For example, in measuring volume with a burette, the uncertainty involved is calculated as follows.

Let the volume delivered = V ml

V = Final burette reading – Initial burette reading

If the uncertainty in both the reading is ±0.02, then the uncertainty involved in the volume delivered in the burette, ΔV = ±(uncertainty in final burette reading + uncertainty in initial burette reading)

i.e. $\pm(0.02 + 0.02) = \pm0.04$

The uncertainty involved in measuring the temperature from the following data is calculated as follows:

5l.2 ± 0.2ºC – 2S.4 ± 0.2ºC

Here a = 5l.2, b = 2S.4

$c = a–b$

= 5l.2 – 2S.4

= 27.8°C

Δa = ± 0.2ºC and Δb = 0.2ºC

Hence, $\Delta c = \pm(\Delta a + \Delta b)$

$= \pm(0.2 + 0.2)$

$= \pm 0.4$ºC

Hence, the net result of the above data is 27.8 ± 0.4.

Uncertainty Involved in Multiplication and Division

If the measured values are multiplied or divided to get the final result, then relative uncertaintiesare added to get the relative uncertainty in the final result.

Consider the following relationship of a multiplication

$c = a$ x b where a and b are values of two measurable quantity. If Δa and Δb are the uncertainty associated with the measurable quantity a and b respectively, then the actual value are $a \pm \Delta a$ and $b \pm \Delta b$ respectively. Therefore, the resulting value will be $c \pm \Delta c$. These are related as follows:

$\Delta c/c = \pm (\Delta a/a + \Delta b/b)$The same relationship holds good for division also.

$c = a/b$ then $\Delta c/c = \pm (\Delta a/a + \Delta b/b)$

For example, when 250 gm of water is heated through l0ºC and if the uncertainty in weight is ±l gm and that in temperature is ±0.2ºC, then the uncertainty involved in the heat absorbed is as follows:

Here a = 250 gm and b = l0°C

Δa = ±l gm and Δb = ±0.2°C

Heat absorbed = mass of water in gram x specific heat of water x temperature to which heatedSpecific heat of water = l

so heat absorbed $c = a$ x b = 250 x l0 = 2500$(\Delta c/c) = \pm(\Delta a/a) + (\Delta b/b)$

$= \pm(l/250 + 0.2/l0)$

$= \pm(0.004 + 0.02)$

$= \pm 0.024$

$\Delta c = \pm 0.024 \times c$

$= 0.024 \times 2500 = 60$

Hence the heat absorbed should be 2500 ± 60 calorie.

2.5 Accuracy and precision

The correctness and reproducibility of a measurement can be expressed in terms of accuracy and precision respectively as follows.

2.5.1 Accuracy

Accuracy is a measure of how closely the result of an experiment agrees with the true or accepted value. In other words, it expresses the correctness of a measurement. However, accuracy is never known. It is known within certain limits only. It can be approached but never be attained. It is because, the results cannot be expressed by any finite number of digits due to mistake made byexperimenter and by the use of measuring device. However, the only kind of physical quantitiesthat can be measured with perfect accuracy are a tally of discrete objects like rupee and coin; etc.

There are two possible ways of determining the accuracy

Absolute method

Comparative method

Absolute method

In this method to determine the accuracy the experiments are repeated several times. This is know as replicate analysis. The consistency of the result in replicate analysis is very often taken as a test of accuracy. However, this is not always so since for some unidentifiable reasons the magnitude of error in every measurement might have been the same and the result consequently might have been consistent although necessarily accurate.

The difference between the mean of an adequate number of results and the actual result may be taken as the measure of the accuracy of the method.

Comparative method

In this method the accuracy is judged by using two or more independent techniques such as gravimetric, titrimetric, spectrophotometric, etc. to solve the same problem. Independent techniques are methods based on different physical and chemical principles. If two independent techniques give the same result the result is thought to be free from error, hence accurate.

In strict sense accuracy can never be determined unless the true value of the measured propertyis known but the limit of the accuracy can be estimated. It is the task of the analyst to judge the best value from

replicate measurements for a property of the given sample. Mathematical and statistical methods are employed determine or to judge the best value and reliability of result in expressingthe accuracy as follows:

2.5.2 Methods of Expressing Accuracy

The accuracy of a measurement can be expressed in terms of

Absolute error,

Relative error and

Relative accuracy.

Absolute error

The difference between the measured value x_m and the true value x_t with regard to sign is called absolute error (X). It is reported in the same unit.

Absolute error, $X = (x_m - x_t)$

Thus, if 2.68 g of a sample of material is analyzed to be 2.59 g, the absolute error = (2.59 – 2.68) g = –0.09 g.

Relative error

The relative error gives the error relative to the size of the measured value. It can be expressed as the percentage or parts per thousand (ppt) of the true value as given below.

Absolute error 100

Percent relative error,

$X_{r\%}$ =_____

2.5.3 **Accepted value or true value**

Parts per thousand error =

The relative error in parts per thousand, $E_{r(0\ 00)}$ = –SS

Relative error can be minimized by precise measurement and eliminating the error of procedure.

Relative accuracy

It is the measured value expressed as the percentage of the true value.Thus When a sample is analyzed several times, the individual results are rarely the same. Instead, the results are randomly scattered. Precision is a measure of how closely the result of an experiment agrees with those of other measurements made in the same ways. In other words, it expresses the reproducibility of the results. This can be achieved unlike accuracy.

2.5.4 Comparison between Accuracy and Precision

It is to be noted that precision always accompanies accuracy but a good precision does not meangood accuracy due to the following reasons.

The same mistake may be made over and over again.

An experimental procedure may be precise but due to some constant and unknown errors the results may be inaccurate.

This can be best illustrated by the following example.

A substance know to contains 42.l0 ± 0.02% of a constituent X. The result obtained by twoanalysts using the same substance and the same analytical method were as follows:

Analyst l% X: 42.0l, 42.25, 42.08, 42.l4, the arithmetic mean is 42.l2% and the result rangesfrom 42.0l% to 42.25%

Analyst 2% X: 42.40, 42.44, 42.42, 42.42, the arithmetic mean is 42.42% and the result ranges from 42.40% to 42.44%. The result of the analysis can be summarized as follows:

The values obtained by analyst l are accurate (very close to the correct result), but precision is inferior to the results given by analyst 2. The values obtained by analyst2 are very precise but not accurate.

The results of analyst l are both sides of the mean values and this may be attributed to random error. However, there isa constant systematic error present in the result of analyst 2.

The various cases of accuracy and precision are shown in Figure 4.2.

The goal in any measurement should be to obtain both precision and accuracy.

2.5.5 Methods of Expressing Precision

While analyzing a sample, each set of analytical results should be accompanied by an indication of the precision of the analysis. The precision can be expressed in terms of mean and median, range or spread, average deviation, relative average deviation, standard deviation, relative standard deviation, variance, standard deviation of the mean, relative standard deviation of the mean and confidence limit as described below.

2.6 The mean and medians

Quite obviously single measurement cannot give an indication of the reliability of results. It is the general practice therefore to carry out replicate analysis using the same method and equipment and obtained a number of values for the same property of a given substance. For example, in titrimetry, an experiment is repeated three or four times with the same solution and the same apparatus in the selection of the best values from the various values. Two methods are commonly adopted to determine average and the median.

The Mean: The term mean, arithmetic mean and average x are synonymous for the numerical value obtained by diving the sum of the values of a set of replicate measurement by the number of measurements made.

where x_i represents the value of ith measurement and n is the number of independent measurements.

The Median: The median (x_{med}) is the value for a set of ordered data for which one half being numerically smaller and the other half being numerically greater. At first the values of the measurement are to be arranged in increasing order. If the set consists of an odd number of measurements (suppose n), the selection of the median may be done directly taking the value of

= l0Exprerrion of precirion by range

The Range: The range (w) of a set of data is simply the difference between the maximum and minimum value in the data set. The range is also called spread.

Range = $w = x_{largest} - x_{smallest}$

Exprerrion of precirion by mean deviation or average deviation, d and relative averagedeviation

It may be defined as the mean of the difference of individual measured value and the mean of the measurement without regard to sign. If $x_1, x_2, \ldots, x_n$ are the values for lst, 2nd, …, nth measurement respectively and x is the mean value, then average deviation where x_i = individual measured values

x = mean of the measurement

Σ represents summation

2.7 Expression of precision by standard deviation, relative standard deviation and variance

Standard deviation (σ or s): The standard deviation σ of an infinite set of experimental data is theoretically the square root of the mean of square of the difference between the individual measuredvalue (x_i) and the ***mean (μ) of the infinite number of measurement.***

In practice, it is only possible to calculate the individual deviation from the mean x of a limited number of measurements, i.e. x_i x. Hence it is desirable to define a quantity which is an experimentally reliable estimate of the standard deviation. Such a quantity is called the estimatedstandard(s), which is applicable to a finite set of data and is given by where n is the number of measurement and the number $(n - 1)$ is called the number of degrees of freedom or independent measurements.

The term degrees of freedom may be defined as the number of individual observations that may be allowed to vary, provided that x and s once determined are held constant. For example, oncethe mean is obtained and we decide to keep it constant then all but one observation can be varied; the last one is fixed by x and all its x values, so there can be $(n - 1)$ number of independent measurement possible, i.e. the degree of freedom = $n - 1$ for n measurements. When n is greater than S0, it is safe to assume $s \rightarrow \sigma$.

Significance of standard deviation: Standard deviation of a set of experimental measurements is a predication that 68 percent of an infinite number of replicate measurements will lie in the interval about the mean. Thus, if in a given case the standard deviation for a set of result with a mean of 12.86 is 0.02, then it means that if an infinite number of measurement are made 68 percent of the measurement will lie in the interval 12.86 ± 0.02 about the mean.

Relative standard deviation: The standard deviation may be expressed relative to mean as per hundred % or per thousand (‰)

Relative standard deviation (%) = Relative standard deviation (‰) =s

Variance, s^2:

Standard deviation of the mean: The standard deviation of the mean is an estimate of the probable error in the mean of a series of observation and is referred to as standard error defined as follows:

Standard error = s (mean) = $\frac{s}{\sqrt{n}}$

where n is the number of measurement and s is the standard deviation of standard deviation.

Relative standard deviation of the mean: Like the standard deviation of the mean, it is possibleto define relative standard deviation of the mean (relative s mean)

The confidence limit: In quantitative analytical work, one deals with relatively small number of measurements. When the number of measurement is a small finite number, one deals with s instead of σ and x instead of μ. s and x are the only estimate of σ and μ. Thus we may conclude that though standard deviation of a set of data

provides an indication of precision inherent in a particularprocedure of analysis, it does not give any information about how close the experimentally determined mean might be with the true mean value μ. Statistical theory allows us to estimate the range within which the true value might fall within a given probability defined by the experimental mean and standard deviation. This range is calledthe confidence interval and the limits of the range are called the confidence limit. The likelihood that the true value falls within this range is called probability or confidence level usually expressed in terms of percent. The confidence limit is given by

Confidence limit = $x - \frac{ts}{\sqrt{n}}$

where t is a statistical factor that depends on the number of degrees of freedom v, (v = n – l) and the confidence of desired level. The values of t at different confidence levels and degrees of freedomare given in Table 4.l. From Table 4.l it is seen that on increasing the number of replicate measurements both the

Table 4.1 t Values for various confidence levels

Number observation (n)	**Number of degree of freedom $n-1$**	**Probability levels 90%**	**9Z%**	**99%**
2	1	6.S14	12.706	6S.660
S	2	2.920	4.S0S	9.925
4	S	2.S5S	S.182	5.841
5	4	2.1S2	2.776	4.604
6	5	2.015	2.571	4.0S2
7	6	1.94S	2.447	S.707
8	7	1.895	2.S65	S.500
9	8	1.860	2.S06	S.S55
10	9	1.8SS	2.262	S.250
11	10	1.812	2.228	S.169
12	11	1.800	2.200	S.110

2.8 Comparing a Mean Value with a True Value (The Student's *t* Test)

W. S. Gosset, an English chemist writing under the pen name of student proposed a test to know whether there is a significant difference between a new method and a standard method or not. This test is called student²s t test as described below.

Two sets of replicate measurement are made by two different methods, one is the new method and the other is the standard method, the two ways in which *t* test can be used are:

A series of replicate analysis are done in a single sample (having the same concentrationby two methods).

A series of analysis are done on a set of different samples (with different concentrationsby two methods).

The student²s *t* value is calculated by applying the following equationFor the first method (a)

$$t \quad x \qquad\qquad |- \quad |\frac{\sqrt{n}}{s}$$

where μ is the true mean value, *s* is the standard deviation, x is the average value for *n* number of observations.

For the second method (b), the difference (D_i) between each of the paired measurement on each sample is computed with regard to sign, i.e. actual sign (±) of the difference is considered and the average difference (D) is calculated and the individual difference each from D,

The calculated *t* value is compared with the tabulated *t* value for a given number of measurements at the desired confidence level (Table 4.l). If the calculated *t* value exceeds the tabulated value then there is a significant difference between the results of the twomethods at that confidence level. If it does not exceed the tabulated value, then we canpredict that there is no significant difference between the two methods.

Illrstration 5 For example, if mean of l2 determination, x = 8.S7 and the true value, μ = 7.9l and standard deviation, *s* = 0.l7, then whether the result is significant or not at 90% confidence level can be decided as follows:

For *n* = l2, the number of degrees of freedom = $12 - 1 = 11$.

From *t* table, for eleven degrees of freedom, the value of *t* at 90% confidence level is l.80. Therefore, the calculated value exceeds the tabulated *t* value. Hence there is a significant differencebetween the results of the two methods at 90% confidence level.

Illrstration 6 Following are the two sets of results for a number of individual samples by a new analytical method and a known standard method to determine whether there is significant difference between the two methods at 95% confidence level.

Sample	New analytical method	Standard method	D_i	$D_i \quad D$	$(D_i \quad D)^2$
A	10.2	10.5	–0.S	–0.6	0.S6
B	12.7	11.9	0.8	0.5	0.25
C	8.6	8.7	–0.1	0.4	0.16
D	17.5	16.9	0.6	0.S	0.09

E	11.2	10.9	0.S	0.0	0.00
F	11.5	11.1	0.4	0.1	0.01
			1.7		0.87

Here, there are six samples for analysis, hence $n = 6$ and number of degrees of freedom $=6 — \mathbf{l} = 5$ The tabulated t value at 95% confidence level for 5 degrees of freedom of is 2.57. Therefore, $t_{calc} < t_{table}$ and there is no significant difference between the two methods at 95% confidence level.

2.9 Comparing Two Experimental Means

Suppose that a sample has been analyzed by two different methods, yielding means x_1 and x_2 and standard deviation s_l and s_2. Let n_l and n_2 are the number of individual results obtained by the two methods. The first step is to calculate a t value using the formula.

$$t \qquad \frac{|\bar{x}_1 \; x_2|}{s} \sqrt{\frac{n_1 n_2}{n_1 \; n_2}}$$

This procedure assumes that s_l and s_2 are the same (nearly same). The second step involve entering t table at a degree of freedom given by $(n_l + n_2 - 2)$ and at the desired probability (or confidence)

2.10 Summary of the Unit

2.11 Key words

2.11 Questions for self Understanding

Short Answer Type Questions

1. Ezplain the followings
2. Significant figure.

3. Leading zero, captive zero and trailing zero giving suitable example.
4. Average deviation.
5. Co-efficient of variation and variance.
6. Confidence interval.
7. Constant error.
8. What do you mean by
9. Degrees of freedom?
10. Determinate error?
11. Gaussian (normal) distribution?
12. Proportional error?
13. Standard deviation?
14. Relative average deviation?
15. Indeterminate error?

Answer the followings

1. Explain clearly the meaning of confidence interval of the mean at(i) 95% and (ii) 99%.
2. Explain clearly how to test two sets of results to determine whether they differ significantlyor not.
3. Ezplain the difference between the followings
4. Accuracy and precision.
5. Random and systematic error.
6. Mean and median.
7. Absolute and relative error.
8. Variance and standard deviation.

Long Answer Type Questions

1. Explain significant figures giving suitable examples. What rules are to be followed for determining the significant figures?
2. Define errors. Mention their causes. Differentiate between determinate and indeterminateerror.
3. Write propagation of errors in addition, subtraction, multiplication and division by giving suitable examples.
4. Define the term accuracy. Why no measurement can be done with absolute certainty?
5. How is accuracy expressed?
6. Define precision. How will you justify the statement that good precision does not assure good

accuracy? What are the different ways of expressing precision?

7. Name and explain different tests of significance.
8. Explain different tests for rejection of data.

Unit-3

Significant figures and rules

3.0 Objectives of the Unit

After studying this Unit ou are able to

3.1 Significant figures

To represent the analytical data obtained from the measurement, certain figures which involve the number of digits are necessary. The figures used to estimate the uncertainty involved in the measurement are called significant figures.

3.2 Definition of Significant Figure

Significant figure in a number may be defined as all the certain digits plus one doubtful (or uncertain) digit. For instance, if the volume of the liquid is measured by a burette, which is graduated, to 0.l ml any volume reported should reflect this. For example, volumes of l2.6 ml and l2.60 ml are numerically the same. But from the analytic point of view l2.6 ml containing three figures (l, 2, 6) is significant when the burette is graduated to 0.l ml whereas actual volume is l2.6 ± 0.l, l2.60 ml containing four figures (l, 2, 6, 0) is significant only if the burette is graduated to 0.0l ml In the above example, the numbers l and 2 are known with certainty while there is some uncertainty in the last digit so that the actual volume could be anything between l2.5 and l2.7 ml. On the other hand, a reading of l2.60 ml implies that the actual volume could be anything between l2.59 and l2.6l when the burette is graduated to 0.0lml, i.e. actual volume is l2.60 ± 0.0l.

Suppose the weight of an object weighed in an ordinary triple beam balance is S.45 g. If the reproducibility of the balance is ± 0.02 g, the second figure past the decimal is uncertain. The weight of the object should lie in the range of S.4S to S.47 g and any effort to report the weight in the third decimal place such as S.450 g by weighing on the above balance is wasted effort. Thus the significant figures in this example are three (S, 4 and 5).

If the same object is weighed in an analytical balance that has a reproducibility of ± 0.002, then the weight of the object should lie in the range of S.448 to S.452 g and the figures (S, 4, 5, 0) in

S.450 are said to be significant. The following rules are helpful in determining the number of significant figures in a given number.

Rules for Determining Significant Figures

3.3 Non- ero integerr

All the non-zero integers are significant. For example, 5.2Sl has four significant figures. 2.45 has three significant figures and 6.25l have four significant figures.

3.4 Recognition of error significant figurer

The digit zero may or may not be significant figure. In the burette graduated to 0.0l ml, let the burette reading be 20.05 ml. This contains four significant figures. Here both zeros are treated as significant figures. If the above burette reading is expressed in litre, it could be 0.02005 litre. Since the number of significant figures can not change just by changing the unit, this number should have only four significant figures. The function of zero preceding 2 (or initial zero) is only to locate the decimal point. Hence initial zeros are not significant similarly a volume of 2 liters expressed in ml. (2000 ml must contain only one significant figure as which is 2. To avoid ambiguity, it is safe to express the number in powers of ten. Thus 2 litres should be written as 2 x l0S ml (one significant figure). If, however, the volume expressed in litre is 2.l, the volume in millilitre will be 2.l x l0S (two significant figures in both cases).

The following examples illustrate the points mentioned

Number	Expression	Significant figures	Remarks
0.612	612×10^{-3}	3	0 is not significant here
70.9	709×10^{-1}	3	0 is significant figure here
0.007	7×10^{-3}	1	Zeros are not significant figures here
	6.023×10^{23}	4	Zero is significant figure here
	1.03×10^{-13}	3	Zero is significant figure here

From the above discussion it is clear that there are three classes of zeros.

Leading xeros: Zeros that precede all the non-zero digits are called leading zeros. These

are not counted as significant figures. They merely indicate the position of decimal point. For example, in 0.007, the zeros are leading zeros so that 0.007 has one significant figure.

Captive xeros: Zeros between two non-zero digits are called captive zeros. These are always counted as significant figures. For example, in 70.9, 0 is captive zero and thus it is taken as significant figure. Therefore 70.9 has three significant figures. Similarly, 5.02 has three significant figures.

Trailing xeros: Zeros at the right end of the number are called trailing zeros. They are significant if the number contains a decimal point. Thus in 900.00, the zeros are example of trailing zero and are significant as these are present after the decimal point. When the number ends in zeros that are not to the right of the decimal point, the zeros are not necessarily significant. For example, the number 240 can have two or three significant figures. Similarly, 20500 can have three, four or five significant figures.

Exponential notation

In exponential notation, the numerical portion represents the number of significant figures. For example, S x l06 has one significant figure S.2 x l0–4 has two significant figures.

3.5 Rounding off the non-significant figurer

The non-significant figures in the measurement are rounded off as per the following rules. If the digit following the last significant figure is greater than 5, the number is rounded off to the next higher digit. For example, 9.48 can be rounded off to 9.5. If it is less than 5, the number is rounded off to the present value of the last significant figure. For example, 9.42 is rounded off to 9.4. If the last digit is 5, the number is rounded off to the nearest even digit. For example,

8.65 is rounded off to 8.6, 8.75 is rounded off to 8.8, 8.55 can be rounded off to 8.6.

The following rules are to be followed while doing calculations involving significant figures.

3.6 Rules for deciding the number of significant figures in a measured quantity

(1) All nonzero digits are significant:

1.234 g has 4 significant figures,

1.2 g has 2 significant figures.

(2) Zeroes between nonzero digits are significant:

1002 kg has 4 significant figures,

3.07 mL has 3 significant figures.

(3) Zeroes to the left of the first nonzero digits are not significant; such zeroes merely indicate the position of the decimal point:

0.001^{o} C has only 1 significant figure,

0.012 g has 2 significant figures.

(4) Zeroes to the right of a decimal point in a number are significant:

0.023 mL has 2 significant figures,

0.200 g has 3 significant figures.

(5) When a number ends in zeroes that are not to the right of a decimal point, the zeroes are not necessarily significant:

190 miles may be 2 or 3 significant figures, 50,600 calories may be 3, 4, or 5 significant figures. The

potential ambiguity in the last rule can be avoided by the use of standard exponential, or ”scientific,”

notation.

For example, depending on whether 3, 4, or 5 significant figures is correct, we could write

50,6000 calories as:

5.06×10^4 calories (3 significant figures)

5.060×10^4 calories (4 significant figures), or

5.0600×10^4 calories (5 significant figures).

3.7 Exact number

Some numbers are exact because they are known with complete certainty. Most exact numbers are integers: exactly 12 inches are in a foot, there might be exactly 23 students in a class. Exact numbers are often found as conversion factors or as counts of objects. Exact numbers can be considered to have an infinite number of significant figures. Thus, number of apparent significant figures in any exact number can be ignored as a limiting factor in determining the number of significant figures in the result of a calculation.

3.8 Rules for mathematical operations

In carrying out calculations, the general rule is that the accuracy of a calculated result is limited by the least accurate measurement involved in the calculation.

(1) In addition and subtraction, the result is rounded off to the last common digit occurring furthest to the right in all components. For example, 100 (assume 3 significant figures) + 23.643 (5 significant figures) = 123.643, which should be rounded to 124 (3 significant figures).

(2) In multiplication and division, the result should be rounded off so as to have the same number of significant figures as in the component with the least number of significant figures. For example, 3.0

(2 significant figures) 12.60 (4 significant figures) = 37.8000 which should be rounded off to 38 (2 significant figures).

3.10 Rules for rounding off numbers

(1) If the digit to be dropped is greater than 5, the last retained digit is increased by one. For example, 12.6 is rounded to 13.

(2) If the digit to be dropped is less than 5, the last remaining digit is left as it is. For example, 12.4 is rounded to 12.

(3) If the digit to be dropped is 5, and if any digit following it is not zero, the last remaining digit is increased by one. For example,

12.51 is rounded to 13.

(4) If the digit to be dropped is 5 and is followed only by zeroes, the last remaining digit is increased by one if it is odd, but left as it is if even. For example,

11.5 is rounded to 12,

12.5 is rounded to 12.

This rule means that if the digit to be dropped is 5 followed only by zeroes, the result is always rounded to the even digit. The rationale is to avoid bias in rounding: half of the time we round up, half the time we round down.

3.11 General guidelines for using calculators

When using a calculator, if you work the entirety of a long calculation without writing down any intermediate results, you may not be able to tell if a error is made and, even if you realize that one has occurred, you may not be able to tell where the error is.

In a long calculation involving mixed operations, carry as many digits as possible through the entire set of calculations and then round the final result appropriately. For example,

(5.00 / 1.235) + 3.000 + (6.35 / 4.0)=4.04858... + 3.000 + 1.5875=8.630829...

The first division should result in 3 significant figures; the last division should result in 2 significant figures; the three numbers added together should result in a number that is rounded off to the last common significant digit occurring furthest to the right (which in this case means the final result should be rounded with 1 digit after the decimal). The correct rounded final result should be 8.6. This final result has been limited by the accuracy in the last division.

Warning: carrying all digits through to the final result before rounding is critical for many mathematical operations in statistics. Rounding intermediate results when calculating sums of squares can seriously compromise the accuracy of the result.

3.12 Ruler for calculations involving significant figurer

Rule for addition and subtraction: The rule here is to retain as many decimals in the final result as the number with the fewest decimal. Let us take the following operation.

l4.22 + 8.l45 – 3.6750 + l20.4 = l39.09

The number, in the above, containing the fewest decimals is l20.4. So the result should contain one decimal. Hence the result l39.09 to be rounded off to l39.l so that it contains one decimal.

Rule for multiplication and division: The rule followed is to retain only as many significant figures as those in the number with the fewest significant figures. For example,

$$\frac{25 \text{ X } 0.524}{100.0} = 0.131\,25$$

In this example, the number 25 has two minimum number of significant figures. Hence, the result should contain only two significant figures. Hence the result 0.l3l is rounded off to 0.l3.

3.13 Rule for logarithm and antilogarithm:

A logarithm is composed of two parts such as a whole number called the characteristic and a decimal fraction called the mantissa. The characteristic is a function of position of the decimal in the number whose logarithm is being determined and therefore, is not a significant figure. The mantissa is the same regardless of the position of the decimal and all the digits are considered significant. For example, consider the logarithm of 2.S x l0S. The characteristic is S. Using a log table, the mantissa is found to be 0.S6l7. Since the number 2.S has two significant figures, its mantissa should contain only the same number of significant figures. Hence mantissa can be expressed as 0.S6. While taking antilogarithm, the result should contain the same number of significant figures as that is mantissa. Thus antilog of l.946 is 99.S. As its mantissa, i.e. 0.946 contains three significant figures, its antilog should also contain the same number of significant figure, that is, three.

3.14 Confidence limit

In chemical analysis, the true value of the population mean μ can not be determined because a very large number of measurements (approaching infinite number) would be required to calculate it. However, with the help of statistics a range can be established surrounding sample mean x within which the population mean μ is expected to lie with a certain degree of confidence based on probability distribution. Thus, “the range (or a numerical interval) around the mean x of a set of replicate analytical results within which the population mean μ can be expected to lie, with a certain degree of confidence (i.e., with a certain probability), is known as Confidence Interval. “The boundaries of this range are called the Confidence Limits. “The likelihood that the true value falls within the range is called the Probability or Confidence Level and it is often expressed as a percentage” consider the example that in a set of iron determination it is 95% probable that the population mean μ lies in the limit ± 0.20% Fe of

sample mean x = 11.30% Fe. It tells that the confidence interval is 11.10% to 11.50% Fe with 95% probability. The 95% confidence limit for μ = 11.30% ±0.20% Fe, and the confidence level is 95%.

The confidence interval is related to the standard deviation of the mean and its size depends on how well the sample standard deviation s estimates the population standard deviation σ. If s is a closer estimate of σ, the narrower is the confidence interval. To give confidence interval at the high confidence levels is one of the best ways of indicating reliability. We shall discuss below two cases for finding the confidence interval: (A) when σ is known or s is a good approximation of σ, and (B) when σ is not known.

3.15 Confidence Interval When σ is known or s is a Good Estimate of σ

We have learned that the normal error curve can be expressed by an Eq. (3.24) in a single variable y where y is defined as ± y = (x – μ)/ σ given by Eq. (3.22). From the definition of y it follows that (x – μ) = ± y σ with the help of normal error curve and the values of y with probability given in Table 3.1 we see that the areas represent probabilities for the absolute deviation | x – μ | to exceed the value of y σ. Since y σ is the deviation of a single observation x from the population mean μ, we can express the probability in terms of y as,

$$\pm y\,\sigma = x - \mu$$

$$\text{Or } \mu = x \pm y\,\sigma \ldots (3.26)$$

Again from Table 3.1, when y = 0.67 there are 50% chances that an observation will

lie in the area having a lower deviation than ± .67 σ, & when y = 1.00 there is a 68.3% probability that a particular measured value has a deviation ± σ and so on. Thus, based on measuring a single value we can write for population mean μ (from Eq. 3.26) lying within limits.

μ = 0.67 σ with 50% confidence (or 50% probability), … (3.27 – i)

μ = x ± 1 σ with 68.3% confidence … (3.27 – ii)

μ = x ± 2 σ with 95.5% confidence … (3.27 – iii)

μ = x ± 3 σ with 99.7% confidence … (3.27 – iv)

Otherwise in more useful way as

$\mu = x \pm 1.96\ \sigma$ with 95% confidence … (3.27 – v)

$\mu = x \pm 2.58\ \sigma$ with99% confidence … (3.27 – vi)

$\mu = x \pm 3.29\ \sigma$ with 99.9% confidence … (3.27 – vii)

However, it is not advisable to estimate the true mean from a single observation x. Instead we generally use the sample mean x to take the better estimate of population mean μ. We also know that for mean x of n observations the standard deviation of the mean is $\frac{1}{\sqrt{n}}$ times of the standard deviation of single observation. Then in terms of $\bar{x}$, the population mean μ lies within the limits.

$\mu = \bar{x} \pm \frac{y\sigma}{\sqrt{n}}$ with the confidence level corresponding to the value y.

That is, confidence interval for $\mu = \bar{x} \pm \frac{y\sigma}{\sqrt{n}}$ or from $\bar{x} - \frac{y\sigma}{\sqrt{n}}$ to $\bar{x} + \frac{y\sigma}{\sqrt{n}}$, and confidence limits can also be expressed with certain confidence level in the same manner.

In the modern practice it is usual to employ a confidence level of 95% for y = 1.96 or

99% for y = 2.58. Thus, we can say that the population mean μ lies within the limits.

$\mu = \bar{x} \pm \frac{y\sigma}{\sqrt{n}}$ with 95% confidence … (3.28 a).

It means that it is 95% probable that the population mean μ lies in the interval $\bar{x} - \frac{1.96\sigma}{\sqrt{n}}$ to $\bar{x} + \frac{1.96\sigma}{\sqrt{n}}$ In a similar way the population mean μ lies within the limits.

$\mu = \bar{x} \pm \frac{2.58\sigma}{\sqrt{n}}$ with 99% confidence … (3.28 b)

$\mu = \bar{x} \pm \frac{3.29\sigma}{\sqrt{n}}$ with 99.9% confidence … (3.28 c)

3.17 Confidence Interval when σ is Not Known

In the usual practice the population standard deviation is not known and can only be

approximated for a finite number of measurements by the sample standard deviation s.

To overcome this difficulty another method is used, which is based on a statistical factor, "t" (also known as student t), that depends on the number of degrees of freedom and confidence level desired. The quantity t is defined as the difference between the two means divided by its standard deviation:

$$\pm t = \frac{\left(\bar{x} - \mu\right)}{s/\sqrt{n}} = \frac{\left(\bar{x} - \mu\right)\sqrt{n}}{s} \quad \ldots (3.29)$$

Statisticians have compiled the values of t for the given degrees of freedom (ν = n – 1) and for various confidence levels desired. For illustration some of the values of t are listed in Table 3.2.

Values of t for υ degrees of freedom and various confidence levels.

υ	Confidence level		
	90%	95%	99%
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.500
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
∞	1.645	1.960	2.576

A simpler and approximate procedure, particularly useful for small numbers of observations (n < 10) to express confidence interval, is based on the range R (the difference between the largest and smallest values of observations). According to this to a certain degree of confidence the population mean μ lies within the limits

$$\mu = \bar{x}, + C_n R \ldots (3.30)$$

where C_n is constant depending upon the number of observations and the desired confidence level. A few values of C are given in Table 3.3.

3.18 Summary of the Unit

Scientific measurements are reported so that every digit is certain except the last, which is estimated. All digits of a measured quantity, including the certain one, are called significant figures. There are conventions that must be followed for expressing numbers so that their significant figures are properly indicated. These conventions are: ALL non-zero numbers (1,2,3,4,5,6,7,8,9) are ALWAYS significant. ALL zeroes between non-zero numbers are ALWAYS significant. ALL zeroes which are SIMULTANEOUSLY to the right of the decimal point AND at the end of the number are ALWAYS significant. ALL zeroes which are to the left of a written decimal point and are in a number >= 10 are ALWAYS significant.

3.19 Key words

3.20 questions for self understanding

1. How many significant figures are in each of the following?

a) 3.405

b) 0.00289

c) 1030

d) 7.0040 x 10-3

e) 102.00

f) 0.000980

g) 9.80

2. Perform the following calculations to the correct number of significant figures

a) 12.0550 + 9.05

b) 257.2 – 19.789

c) (6.21 x 103

) (0.150)

d) 0.0577 ÷ 0.753

e) 27.5 x 1.82 ÷ 100.04

f) (2.290 x 106) ÷ (6.7 x 104)

g) [(28.7 x 105) ÷ 48.533] + 144.99

3. Round each of the following numbers to three significant figures:

a) 342.79513

b) 9,845.8749

c) 0.000045389

d) 2.45555567

e) 76.89

f) 56.9971

Unit-4

Comparison methods

4.0 Objectives of the Unit

After studying this unit you are able to

4.1 Introduction

4.2 Criteria for rejection of data

Sometimes, in a set of analytical data there appears a value which is not fitting in the set as it looks at a wide difference from the rest of the values. Now the question arises how to decide whether to remove out a result which appears out of line with others when there are no known reasons to suspect it? The question is not of much importance if the number of replicate observations is large since a single value will have only a small effect upon the mean. But it is, of course, important when the number of replicate measurements is small, since here the divergent observation has a significant effect on the value of mean, while at the same time there are insufficient data to decide the fate of the suspected result. In a small set of data the decision for rejection or retention by the blind application of statistical test is no doubt an arbitrary decision.

The criteria for rejection of an observation are based on the supposition that an outlier is due to some systematic source of error. If it is not, then it falls within the random error and should be retained. So many authors agree that the question of rejecting one divergent value from a small sample can not satisfactorily be answered. It is unfortunate fact that no universal rule can be invoked to settle the question of retention or rejection. Out of various rules we shall discuss below only two of the more widely recommended criteria for rejection of outlier: one is based on the average deviation and is called the "4d" rule, and the other which is based on the range is called the "Q" test.

4.3 The "4d" Rule

This test is based on the fact that 4 times of average deviation which is an estimate of ($4 \times$ a.d. $= 4 \times 0.80\ \sigma$) $= 3.2\ \sigma$ which lies in the probability distribution curve at a confidence level of

99.8% that a value of measurement lies within this limit. To apply this rule there must be at least 4 observations excluding the outlier (preferably 10 to 30 observations) and the following procedure is adopted:

- ✓ Omit the doubtful value and find out the arithmetic mean of the rest.
- ✓ Find out the average deviation of rest of the values obtained after omitting the doubtful value. Call this as "d" and its four times as "4d" (i.e. 4× a.d.)
- ✓ Obtain the difference between the doubtful value and the mean calculated in (i) call this difference as $z\left(z = \left|xs\text{-}\bar{x}\right|\right)$ where xs is the suspected value and $\bar{x}$ is the mean.
- ✓ Criterion: If the difference z is greater than 4 times of average deviation calculated in (ii) the suspected value should be rejected otherwise retain it. That is, If z > 4d, then the doubtful value is rejected, and if z 4d, then the doubtful value is retained.

4.4 The "Q" Test

Another criterion used to check the rejection of suspected result (in a set of 3 – 10 results) is the Q-test which is a simple and widely used statistical test. Q, the rejection quotient, is defined as the ratio of the divergence of the suspected value from its nearest neighbour to the range of the set of measured values. If the value of Q calculated is greater than the value of Q given in the table at the desired confidence level for the given number of observations, the suspected value is rejected. The Q testis applied as follows:

- ➢ Arrange the observations either in the increasing or decreasing order. The lowest or the highest or both may be the doubtful values.
- ➢ Calculate the range, R = highest values – lowest value
- ➢ Find the difference between the suspected value and its nearest neighbour. Call this difference as Y.
- ➢ Calculate the rejection quotient, Q = Y/R and call it as Qcalculated.
- ➢ Consult the table of Q (Table 3.4) for the given number of observations and call it as Qtabular.
- ➢ Criterion: Compare Q_{calc} with Q_{tab}. If $Q_{calc.} > Q_{tab}$, then the suspected value is rejected.

Note: Since both lowest and highest values may be considered to be the suspected values, it is advisable to apply the test for both. Say in above if the lowest value was taken as the suspected value we have to extend the test as below.

- ➢ If the lowest value is rejected, then the range R' for remaining values is calculated and

if the lowest value is not rejected, then the same range R is used and apply the above procedure steps (ii) to (vi) for the highest value.

- Repeat the process further if necessary. Some Q values are given in Table 3.4.

Table 3.4: Values of Rejection Quotient Q at 90% confidence level

No of observations (n)	Q
3	0.94
4	0.76
5	0.64
6	0.56
7	0.51
8	0.47
9	0.44
10	0.41

The following example will illustrate application of the Q test.

Example Apply Q-test to check the rejection of the highest value in the following results:

2.18, 2.19, 2.30, 2.15 and 2.20

R = 2.30 – 2.15 = 0.15

Y = 2.30 – 2.20 = 0.10

Q_{calc} = Y/R = 0.10/0.15 = 0.67

From Table 3.4, for n = 5, Q_{tab} = 0.64 $Q_{calc} > Q_{tab}$, therefore the value 2.30 is rejected at 90% confidence level.

SAQ 5

In replicate determination of iron the following results of percentage of iron were obtained. Should any of the results be rejected?

%Fe: 52.40, 52.47, 52.50, 52.51, and 52.46.

……...

……...

……...

……...

4.5 Tests of significance

In analytical chemistry we develop new methods of analysis and it is frequently desired to compare the results of a new method with those of an accepted (say from a past experience or a standard, e.g.; from the National Institute of Standards and Technology, NIST) method. The average values obtained from the two methods may show a difference. The question arises whether this difference is due to random fluctuation (indeterminate error) or directional fluctuation (systematic error). The answer is qualified by a degree of certainty involving the method what is known as **null hypothesis** which considers that there is no significant difference between two sets of data. In null hypothesis procedure a comparison, of statistical parameters (based on mean or standard deviation, or some other property), is made between two sets of replicate measurements obtained by two different methods, one of them being the test method and the other usually being accepted method. With this comparison the value of the statistical parameter of test of significance is calculated and compared with the value of the parameter given in the statistical tables available. A simple examination of the two values (calculated & tabular) will show how large a difference needs to be in order to be considered for the limit of significant divergence. Thus, if there is not a statistically significant divergence, means the null hypothesis is valid or that there is no source of systematic error and the variation in results follows the law of random errors. And if there is a statistically significant divergence, means the null hypothesis is not valid and a source of systematic error is highly probable.

We shall discuss below three tests of significance:

(i) the t-test which is based on comparison of two means,

(ii) (ii) the F-test which is based on the comparison of two variances, and

(iii) (iii) the chi^2 – test (χ^2 – test) which is given in terms of frequencies.

4.6 The t-test or Student's t Test

The t-test is used to test the null hypothesis that two means do not differ significantly. The application of t-test will be considered here only in a simple case.

When Accepted Mean Value is Known: Eq. (3.31) used to get the confidence limit is also applicable to the comparison of the finite sample mean *x* and the population mean μ. The quantity t is defined as

$$\pm t = \left(\bar{x} - \mu\right)\sqrt{n}/s \quad \pm\text{----------} \quad (3.31)$$

where $\bar{x}$ = average value for the finite series

μ = population mean when the series has been carried to an infinite number of observations, or accepted value given by some national standards

s = standard deviation of the finite series, and

n = number of measurements in the finite series.

Values of *t* with reference to probability levels of 90, 95 and 99 percent are summarized in Table 3.2. In the table n refers to the degrees of freedom. In the procedure to apply the null hypothesis the quantity $\pm t = \left(\bar{x} - \mu\right)\sqrt{n}/s$ is calculated for the given observations and known as $t_{calculated}$. This value of t is compared with the corresponding value of t (t – tabular) found in table of t (Table 3.2) at the desired confidence level and corresponding to n degrees of freedom of finite sample (ν = n – 1). On comparison,

- If $t_{calc} < t_{tab}$, the null hypothesis is valid, there is no significant difference between the two means (x and μ), the variation in results is just by random errors and no systematic source of error is probable.
- If $t_{calc} > t_{tab}$, then the null hypothesis is incorrect, a significant difference between the two means is indicated and the difference is due to some source of systematic error in the values of finite series.

The above criteria indicate that the smaller the calculated t value, the more confident you are that there is no significant difference between the two means. Suppose five observations obtained for the determination of atomic mass of cadmium were: 112.25, 112.36, 112.32, 112.21, 112.36. Does the mean of these values differ significantly from the NBS accepted value 112.41?

The te

x	$\left\|x-\bar{x}\right\|$	$\left(\left\|x-\bar{x}\right\|\right)^2$
112.5	0.05	0.0025
112.36	0.06	0.0036
112.32	0.02	0.0004
112.21	0.09	0.0081
112.36	0.06	0.0036
$\sum x = 561.5$		Sum=0.0182

st is applied as follows after calculating the requir0.00ed quantities

x = 561.5/5

= 112.30

s = (0.0182) ^{½} = 0.067 and n = 5

$$\pm t = \left(\bar{x} - \mu\right)\sqrt{n}/s$$

= (112.30 – 112.41) (5) ^{½} /0.067

= – 0.11 × 2.236/0.067 = – 3.67

Or t_{calc} = 3.67 (disregarding the negative sign)

From Table 3.2 at 99% probability level corresponding value of t for 4 degrees of freedom is 4.60. Thus t_{tab} = 4.60.

Comparing two *t* values we see that

t_{calc} < t tab at 99% confidence level.

It is concluded that the null hypothesis is valid at 99% probability level and there is no significant difference between the two means. The variation is due to indeterminate errors.

4.6 F-Test

The F test serves to show whether the precision of two different methods is the same within specified probability limits. It is applied in terms of variance ratio. The F value which is the ratio of two variances in question is determined by the relation

$$F = \frac{V_1}{V_2} = \frac{S_1^2}{S_2^2}$$

Placing of the larger of the two variances in numerator $\left(S_1^2 > S_2^2\right)$ so that the value of F is always greater than unity. The value of *F* determined by the use of Eq. (3.51) for experimental variances S_1^2 and S_2^2 is known as $F_{calculated}$.

Statisticians have compiled tables of F values for various significance levels for various degrees of freedom ($\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$), where n1 is the number of observations for the set of larger variance (i.e. larger standard deviation). Table 3.5 is a brief F table (two sided) for the 95% confidence level. The value of F obtained from such a table is called as F tabular.

To test null hypothesis for the two sets of data by F test, the calculated value of F is compared with the corresponding tabular value of F. On comparison

(i) If $F_{calc} < F_{tab}$, that is if the experimental F is smaller than the corresponding tabular value of F, then no statistically significant difference is indicated between s1 and S_2 (i.e. between two sets of data), and the null hypothesis is valid, and

(ii) If $F_{calc} > F_{tab}$, then S_1 is significantly greater than S_2, and the null hypothesis is not valid.

Table 3.5: Values of F at 95% confidence level

	ν_1					
ν_2	2	3	4	5	6	∞
2	19.00	19.16	19.25	19.30	19.33	19.50
3	9.55	9.28	9.12	9.01	8.94	8.53
4	6.94	6.59	6.39	6.26	6.16	5.63
5	5.79	5.41	5.19	5.05	4.95	4.36
6	5.14	4.76	4.53	4.39	4.28	3.67
∞	3.00	2.00	2.37	2.21	2.10	1.00

To illustrate "F" test suppose that two series of observations are made one of 4 observations of standard deviation equal to 0.02 and another of 6 observations of standard deviation equal to 0.04. We have to test whether there is significant difference between the two standard deviations.

For the condition of application of F test we have to consider the greater standard deviation that is, 0.04 as s1 and the smaller 0.02 as s_2. Hence $\nu_1 = 6 - 1 = 5$, and $\nu_2 = 4 - 1 = 3$

F is calculated as

$$F_{calc} = \frac{V_1}{V_2} = \frac{S_1^2}{S_2^2} = \frac{(0.04)^2}{(0.02)^2} = 4.0$$

Corresponding tabular value of F for $u_1 = 5$ and $u_2 = 3$ from Table 3.5 at 95% confidence level is $F_{tab} = 9.01$.

An examination of two F values you find that Fcalc < Ftab, therefore, it is concluded that the null hypothesis is valid and there is no statistically significant difference between the standard deviations of the two sets of data or statistically no significant difference is observed between the precisions of the two sets of data.

4.7 The χ^2 (chi-square) Test

Chi-square test is applied to study the behaviour of data if the theoretical behavior can be expressed quantitatively in terms of expected frequencies. The test is applied to check the number bias (if any) for a particular digit in instrument reading. A number bias varies considerably from observer to observer and also depends on the number ofsub-division of instrument readings. Naturally, for a limited number of observations, a certain fluctuation is statistically expected. However, if the fluctuation, is such that is not governed by probability, can be due to a number bias. Such a number bias can actually impose a limitation upon the accuracy of readings by an individual. The chisquare test informs us about such a number bias. This test gives the comparison of a number of frequency distribution. The quantity chi-square is defined by

c2 = _(*f*i – *F*i)2/*F*i … (3.33)

where *f*i = observed frequency

*F*i = expected frequency

The calculated Chi-square applying Eq. (3.33) is compared with the values of chisquare given in the table for the corresponding degrees of freedom. In Table 3.6 some critical values of chi-square are listed. On comparison of the two,

i) if $x^2_{cal} > x^2_{tab}$, there is a number bias,

ii) if $x^2_{cal} < x^2_{tab}$, there is no number bias, and fluctuation is by chance.

To illustrate let us consider the tossing of a coin for 500 times. The expected results or theoretical results should give expected frequency of 250 times head and 250 times tail. It is rare that these results are obtained exactly. Suppose 270 times we get head. Then 250 is the expected frequency Fi, and 270 is the observed frequency fi. From

Eq. (3.52) we get

χ^2 = (270 – 250)2/250

= $(20)^2$ / 250 = 1.6

Thus, χ^2_{cal} = 1.6

From Table 3.6 for (v = 2 – 1 = 1) one degree of freedom at 95% of confidence level χ^2 tab is 3.84. On comparison we get, $\chi^2_{cal} < \chi^2_{tab}$.

It is concluded that there is no number bias and the fluctuation is by chance. As a second example if we were to examine the last digit of the students burette readings estimated to 0.01 mL, we would find a considerable number bias in favour of the last digits as 0 and 5.

4.8 The least square method

The least square method is the process of finding the best-fitting curve or line of best fit for a set of data points by reducing the sum of the squares of the offsets (residual part) of the points from the curve. During the process of finding the relation between two variables, the trends of outcomes are estimated quantitatively. This process is termed as regression analysis. The method of curve fitting is an approach to regression analysis. This method of fitting equations which approximates the curves to given raw data is the least squares.

It is quite obvious that the fitting of curves for a particular data set are not always unique. Thus, it is required to find a curve having a minimal deviation from all the measured data points. This is known as the best-fitting curve and is found by using the least-squares method.

Linear least squares regression has earned its place as the primary tool for process modeling because of its effectiveness and completeness. A method of determining the curve that best describes the relationship between expected and observed sets of data by minimizing the sums of the squares of deviation between observed and expected values. The regression calculations attempt to minimize this sum of the squares, hence the name “least squares regression

4.9 Least Square Method Definition

The least-squares method is a crucial statistical method that is practised to find a regression line or a best-fit line for the given pattern. This method is described by an equation with specific parameters. The method of least squares is generously used in evaluation and regression. In regression analysis, this method is said to be a standard approach for the approximation of sets of equations having more equations than the number of unknowns.

The method of least squares actually defines the solution for the minimization of the sum of squares of deviations or the errors in the result of each equation. Find the formula for sum of squares of errors, which help to find the variation in observed data.

The least-squares method is often applied in data fitting. The best fit result is assumed to reduce the sum of squared errors or residuals which are stated to be the differences between the observed or experimental value and corresponding fitted value given in the model.

There are two basic categories of least-squares problems:

Ordinary or linear least squares

Nonlinear least squares

These depend upon linearity or nonlinearity of the residuals. The linear problems are often seen in regression analysis in statistics. On the other hand, the non-linear problems are generally used in the iterative method of refinement in which the model is approximated to the linear one with each iteration.

4.11 Least Square Method Formula

The least-square method states that the curve that best fits a given set of observations, is said to be a curve having a minimum sum of the squared residuals (or deviations or errors) from the given data points. Let us assume that the given points of data are (x_1, y_1), (x_2, y_2), (x_3, y_3), …, (xn, yn) in which all x's are independent variables, while all y's are dependent ones. Also, suppose that f(x) is the fitting curve and d represents error or deviation from each given point.

Now, we can write:

$d_1 = y_1 - f(x_1)$

$d_2 = y_2 - f(x_2)$

$d_3 = y_3 - f(x_3)$

…..

$d_n = y_n - f(x_n)$

The least-squares explain that the curve that best fits is represented by the property that the sum of squares of all the deviations from given values must be minimum, i.e:

$$s = \sum_{i=1}^{n} d_i^2$$

$$s = \sum_{i=1}^{n} \left[y_i - f_{x_i} \right]^2$$

$$s = d_1^2 + d_2^2 + d_3^2 + --------+ d_n^2$$

Sum = Minimum Quantity

Suppose when we have to determine the equation of line of best fit for the given data, then we first use the following formula.

The equation of least square line is given by Y = a + bX

Normal equation for 'a':

$\sum Y = na + b\sum X$

Normal equation for 'b':

$\sum XY = a\sum X + b\sum X2$

Solving these two normal equations we can get the required trend line equation.

Thus, we can get the line of best fit with formula $y = ax + b$

4.12 Summary of the Unit

Statistical analysis is necessary to understand the significance of analytical data. In this unit you have studied the methods used by scientists in evaluating the significance of analytical data with the knowledge of normal distribution of errors in terms of probability. Replicate determinations should be made in order to approach the value of experimental mean around the true mean with a certain degree of probability. To indicate the precision of an analysis the most important statistical parameter is the standard deviation. A clearer picture of data quality is sometimes obtained by relative standard deviation. Precision of the calculated results is estimated by the calculation of standard deviation by taking care of uncertainties of different sets of data used in computation of results. You have learnt here that the study of probability distributions is of fundamental importance to the use of statistics for judging the reliability of analytical data. What we say is all measurements contain random errors which follow the normal law of It can be expressed by a differential Eq. The area under various limits of the normal error curve can be calculated by the integration of the differential Eq. The area is the measure of probability that an observation lies in these limits. Various confidence limits are used to get the confidence interval. The rejection of an outlying observation can be ascertained by the suitable statistical methods. In analytical chemistry it is frequently desired to compare the results of two different methods, one of them is the test method and the other is usually an accepted method.

This is done by null hypothesis using the tests of significance: the t-test which is based on the comparison of two means, the F-test which is based on the comparison of two variances, and the chi-square test which is given in terms of frequencies. In many analyses the relation between a physical quantity measured and concentration on plotting gives a straight line. The method of least square is used to obtain the best straight line for which the sum of the squares of deviations of the points from the line is minimum. For quality control, the industrial analytical laboratories often use the control chart technique which tells the trends in data of certain analysis.

4.13 Key words

4.14 Questions for self understanding

1. Consider the following set of replicate measurements of an analyte: 0.792, 0.794, 0.813 and 0.900 g. The true value is 0.830 g. Calculate (a) mean (b) median (c) range (d) standard deviation (e) coefficient of variation (f) absolute error of the mean (g) the relative error of the mean in parts per thousand. Consider no observation is rejected.
2. 2. Calculate the uncertainty of the operation $y = a/b$. The individual uncertainty (as standard deviation) of each quantity is given in parenthesis: $a = 36.2 (\pm 0.4)$; $b = 27.1 (\pm 0.6)$. Express the calculated result with absolute uncertainty.
3. An analyst got the percent alcohol content in a blood sample: 0.084, 0.089 and 0.079. Calculate the 95% confidence limit for the mean assuming $t = \pm 4.30$ for two degrees of freedom and 95% confidence.