Unit 4

Analytical Chemistry

(By : Dr. Shveta Trivedi
Dr. Bhavin V. Patel)

4.1 Introduction :

Analytical chemistry is concerned with the chemical characterization of the matter. Everything we use or consume in our daily life are chemicals and related substances. Analytical chemistry plays an important role in various fields like agricultural chemical, environmental, forensic, industrial manufacturing, metallurgical and pharmaceutical chemistry, pollution control and monitoring etc. ....

Definition : Analytical chemistry deals with the theory and practice of methods used to determine the composition of matter (sample).

Definition : "**Chemical analysis** is defined as the applications of a process or a series of processes for identifying and / or quantifying a material and its components or determination of structures of chemical compounds."

4.2 Qualitative and Quantitative Analysis :

Analytical chemistry can be divided into areas called qualitative and quantitative analysis.

Definition : "**Qualitative analysis** is the process of identification of components of material sample and impurities present". For example Inorganic Qualitative analysis and organic Qualitative analysis carried out in laboratory.

Definition : "**Quantitative analysis** is the process of determining the quantity (amount) of components present in the sample material". For example volumetric analysis and gravimetric analysis carried out in chemistry laboratory.

Further, chemical analysis can be classified on the basis of data to be generated (type of information required) and size (quantity) of sample as :

4.2.1 Classification of Chemical Analysis :

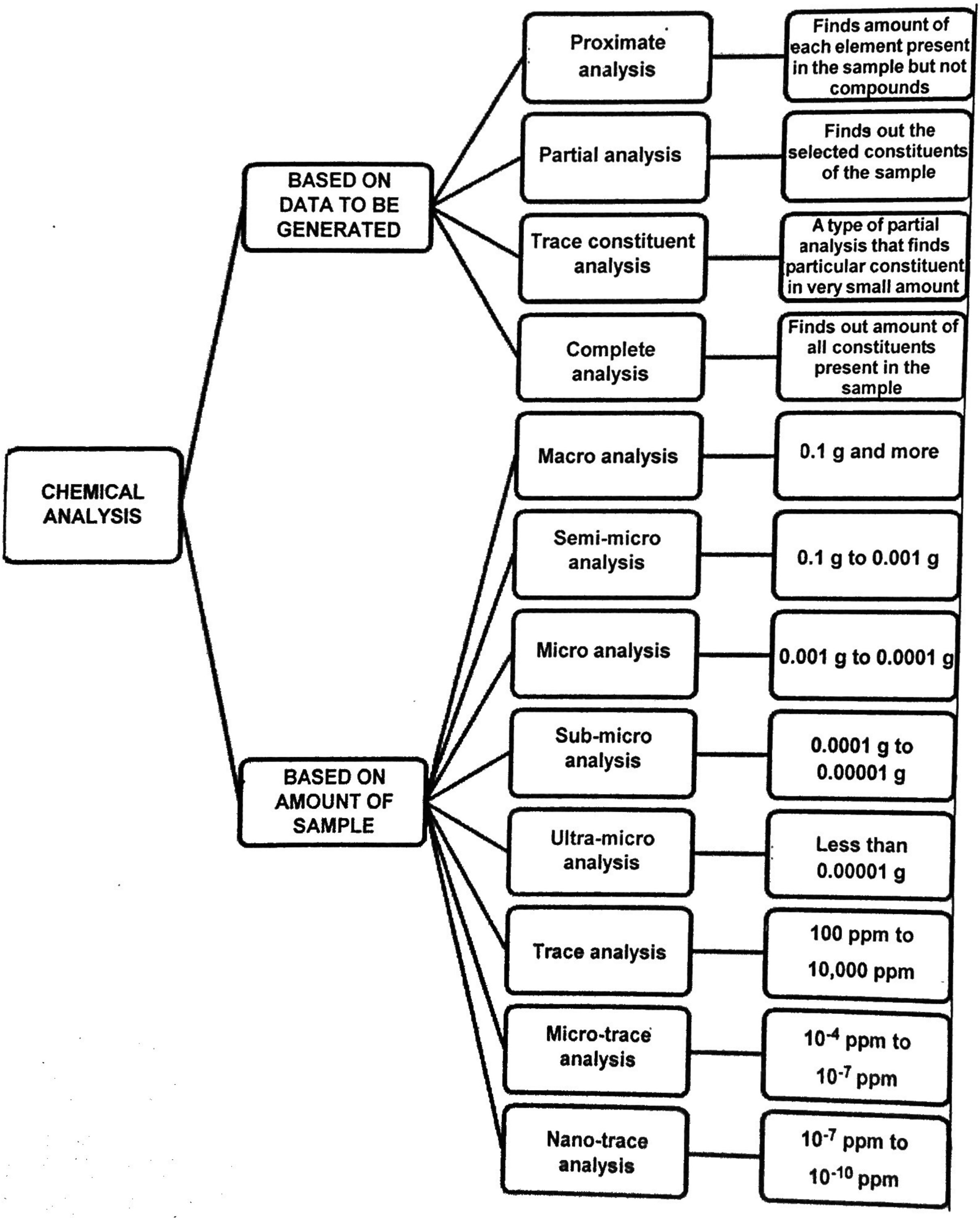


FIGURE 4.1 : CLASSIFICATION OF CHEMICAL ANALYSIS

4.3 Instrumental Methods and Chemical Methods of Analysis :

As discussed earlier, Analytical chemistry can be divided into qualitative and quantitative analysis; it can be further classified into chemical and instrumental methods.

There are numbers of methods under this category.

❆ **Chemical Methods :** "The type of analytical methods where no instruments are used are known as chemical methods. These are comparatively simple to follow but sometimes time consuming." Titration or volumetric analysis, Gravimetry, precipitation are the example of chemical methods. These are also known as non-instrumental methods of analysis.

❆ **Advantages of Chemical Methods :**

(i) Procedure is accurate and simple.
(ii) The equipment needed is cheap/economic.
(iii) Methods are based on absolute measurements.
(iv) Specialised training is usually not required.

❆ **Limitations of Chemical methods :**

(1) Chemical environment is critical.
(2) There is a lack of versatility.
(3) Accuracy decreases with decreasing amounts.
(4) Procedure is time consuming.
(5) There is a lack of specificity.

❆ **Instrumental Methods of Analysis :**

"The type of analytical methods where chemical analysis is/are carried out strictly with instruments are known as instrumental methods of analysis." These methods are quick and more accurate but sometime more expensive. For example pH metery, potentiometer, refractometry, Gas chromolithography, HPLC, TGA, DTA etc. ...

❆ **Advantages of Instrumental Methods :**

(i) Small samples can be used.
(ii) High sensitivity is obtained.
(iii) Measurements obtained are reliable.
(iv) The determination is very fast.
(v) Even complex samples can be handled easily.

● **Limitations of Instrumental Methods :**

(1) An initial or continuous calibration is required.
(2) The sensitivity and accuracy depends on the intrument.

(3) The cost of equipment is high.

(4) The concentration ragne is limited.

(5) Specialised training is needed.

(6) Sizable space is required.

4.4 Applications of Chemical Analysis :

1. Manufacturing industries depend on chemical analysis to ensure the **quality of raw materials, intermediates and the quality of the final product**. The manufactured final product is checked by quality control to ensure that it contains essential components within a required range of composition and the impurities do not exceed certain specified limits.
2. Many industrial processes give rise to pollutants which can create a health problem. **Quantitative analysis of air, water, and in some cases soil samples**, must be carried out to determine the level of pollution and also to establish safe limits for pollutants. Such analysis needs knowledge and proper skill in the analytical field.
3. In hospitals, chemical analysis is widely used to help in the diagnosis of illness and in monitoring the condition of patients by testing urine, blood samples.
4. In farming, the nature and level of fertilizer application is based upon information obtained by analysis of the soil to determine its content of the essential plant nutrients, nitrogen, phosphorus and potassium, and of the trace elements which are necessary for healthy plant growth.
5. Geological surveys require the services of analytical chemists to determine the composition of the numerous rock and soil samples collected in the field.
6. Some special applications like : the metal ions can exist in several forms in natural water. This is because of presence of various substances in water e.g. copper in pure water exists as $Cu(H_2O)_6$ only whereas in natural water, it is presents in the form of different compounds and complexes. Due to difference in properties of each of these species, their biological and geological effects are different.

4.5 Sampling of Solid, Liquid and GAS :

Before carrying the actual analysis of the given sample, the chemist needs to follow certain steps to get that small quantity of analysable sample.

Definition : "**Sampling** involves obtaining a small mass of a material whose composition accurately represents the bulk of of the material given for the analysis".

Sampling must be accurate and must have the same composition of bulk (big mass). generally the bulk is large and heterogeneous. So a great effort is required to get a real sample.

Definition : Real Sample : Real sample is the true representative of the bulk which has same composition as the bulk has. It is obtained from the process known as sampling. Samples can be gas, liquid or solid.

4.5.1 Sampling of Solids :

When our sample is solid in nature, it is generally heterogenous in the composition. When we take samples, everytime it gives different results because of difference in compositions.

After the sample is selected, it is crushed or ground properly to the required particle size and then it is mixed properly so that homogeneity in composition will be obtained. The grinding can be done with mortar and pestle in laboratory or for more bulk, we can use crusher and grinders. Then this material is passed through certain size (mesh) of sieves so that uniform particle size is obtained one of the important method of sampling solid sample is "piling."

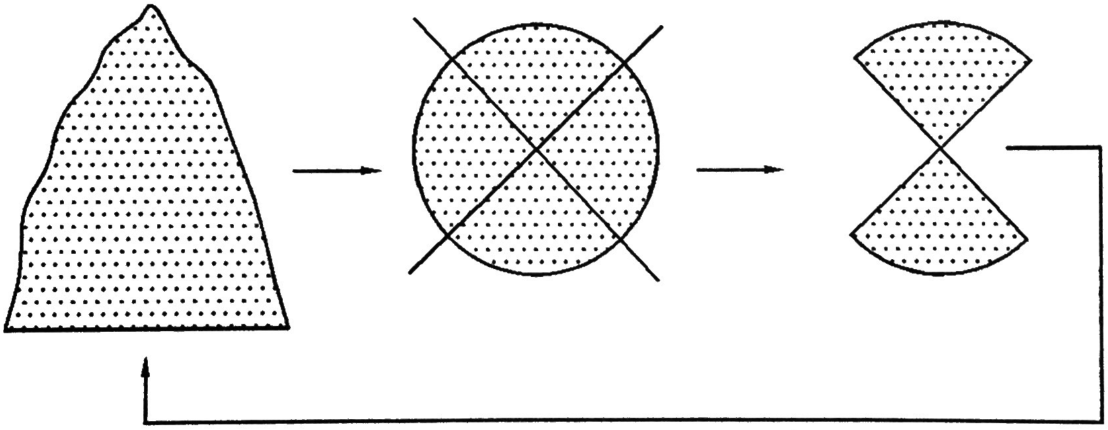


FIGURE 4.2 SAMPLING OF SOLIDS

First of all make a pile of all crushed material and it looks like a cone as shown in the figure. Then make it flat i.e. flattening the cone. After this, divide this into four equal parts and any two opposite parts are collected and remaining two are discarded. The two collected parts are further mixed and again pile is made the new pile is now half in size and this is repeated untill we get the required reduced size (amount) of sample.

4.5.2 Sampling of Liquid :

If our liquid sample is homogeneous, then sampling process is very easy. We can collect the required quantity of sample from anywhere from the bulk. For example : Solutions in Laboratory

But when the liquid sample is heterogeneous, we need to take care to get it homogeneous, chemist can simply shake the liquid sample vigrously so that it becomes homogeneous and the collect it to the required quantity. For example liquid cough syrup. Analyst can also collect the samples from different sites of the bulk so that a true representative can be obtained. Comparatively sampling of liquid is easy and less labour compared to solid samples.

4.5.3 Sampling of Gas :

Because of nature of diffusion of gas, sampling of gaseous sample is comparatively easy. Now a days sampling of gas is more frequent for atmospheric pollution control and monitoring. To check the quality of air we breathe. Air is the complex mixture of different gaseous components and particulate matter. The composition depends upon number of factors such as location, temperature, wind, rain and seasons.

After collecting air samples from various locations, it is passed through various filters through which solid particulates are removed and will not interfere in the actual analysis. Some gaseous impurities can be removed by passing this sample from columns containing certain solutions, which will absab the impurity and can be removed. For example if we pass gaseous sample from a column containg NaOH solution, then it will absorb (remove) CO_2 gas by converting it to Na_2CO_3 precipitates.

4.6 Stages of Analysis

A complete chemical analysis consists of a series of steps as shown in the figure-4.3. In order to **maintain accuracy and reproducibility of results, all these steps must be carried out carefully**. Sample preparation is considered to be the most important stage of analysis. Different methods depending on the nature and size of sample must be used for preparation of the sample. If all the steps are carried out without involvement of humans, the procedure is said to be fully automatic.

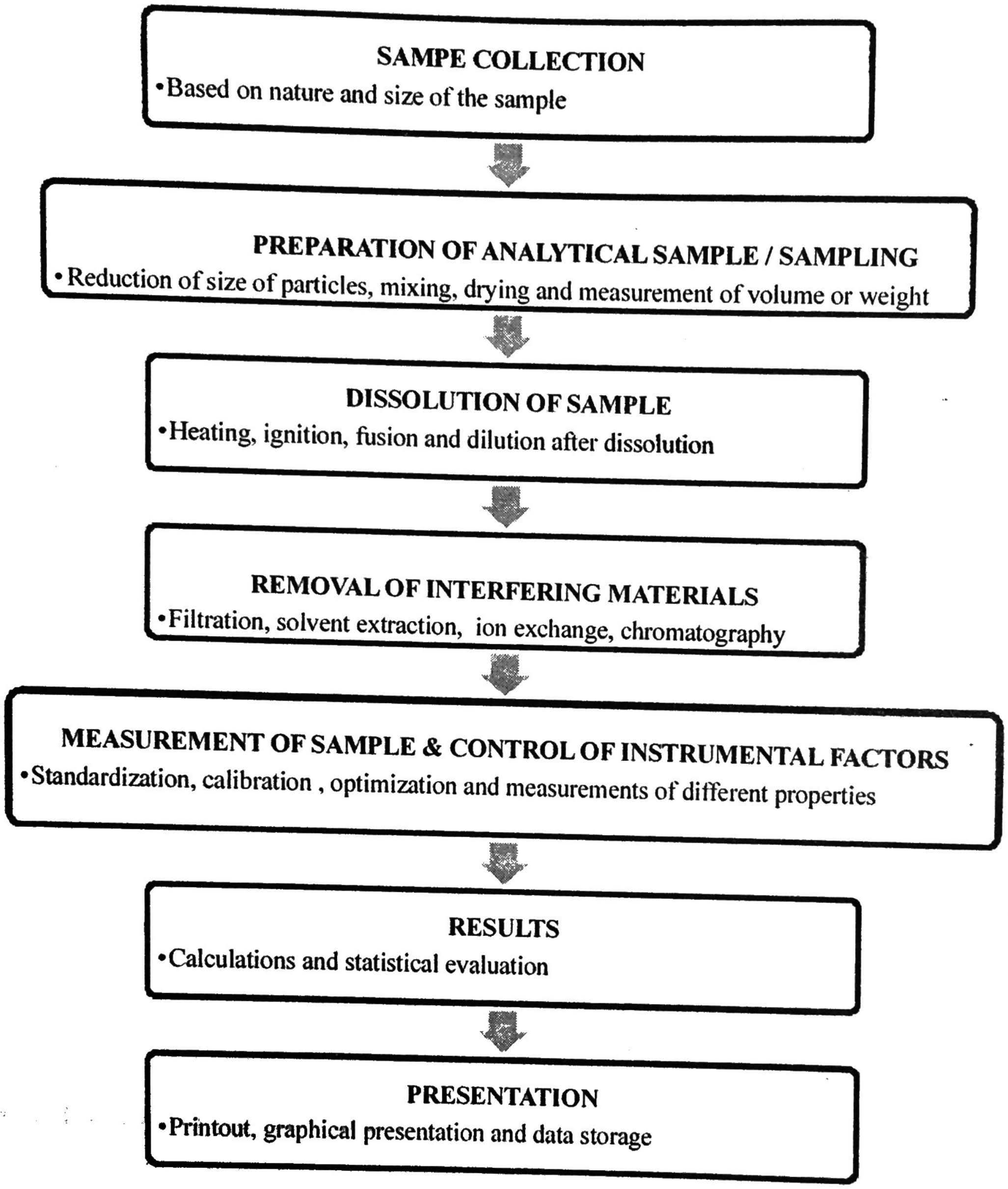


FIGURE 4.3 : STAGES OF ANALYSIS

4.7 Interferences :

Interferences are the substances which are present in the sample and contributes the sample reading. So it is compulsory to remove it before we start our actual sample analysis. There all various prescribed analytical methods for removal of such interferences from the sample. We can select any suitable method according to the nature and type of interference. Some of the frequently used methods are :

(a) **Selective Precipitation** : The interfering ion can be selectively precipitated by knowing the suitable precipitating agent at the given chemical conditions. Sometimes pH control is required for effective removal of interference. A great care is required to ensure that no sample ions are practically lost with this precipitation.

(b) **Masking** : It is a process of trapping an unwanted ion from the solution by adding suitable complexing agent (chelating agent) so that the ion gives precipitates and can be removed easily.

(c) **Selective oxidation or reduction** : Here the interfering ion is either oxidised or reduced in the solution so that during actual analysis it will not take part into reaction and becomes passive eventhough remaining with the sample ions. For example, while precipitating iron as hydroxide it is always Oxidised to Fe (III) State and not Fe (II)

(d) **Solvent Extraction** : It is the treatment of ion solution with immersible organic solvent so that the ion can be extracted by that organic solvent containing suitable chelating (complexing) agent and can be removed.

(e) **Ion Exchange** : This includes utilizing proper cation exchangers or anion exchangers resins for removing unwanted interferences from the solution. These exchanger resings are filled in columns and sample solution is passed through these columns. The resulting solution is free from interferences.

(f) **Chromatography** : It is a separation technique, where by using different stationary phase columns, unwanted ions or species can be removed. There are many methods under this title, some of them are; Adsorption chromatography, partition chromatography, Gas chromolithography, HPLC etc. ....

4.8 Selection of Methods :

The selection of method for analysis mainly depends on the following factors:

1. The type of information to be found.
2. Amount of sample available.
3. Amount of components to be determined.
4. Purpose of analysis.
5. Accuracy required during analysis
6. Possible interferences present in sample
7. Facility available
8. Time available for the analysis

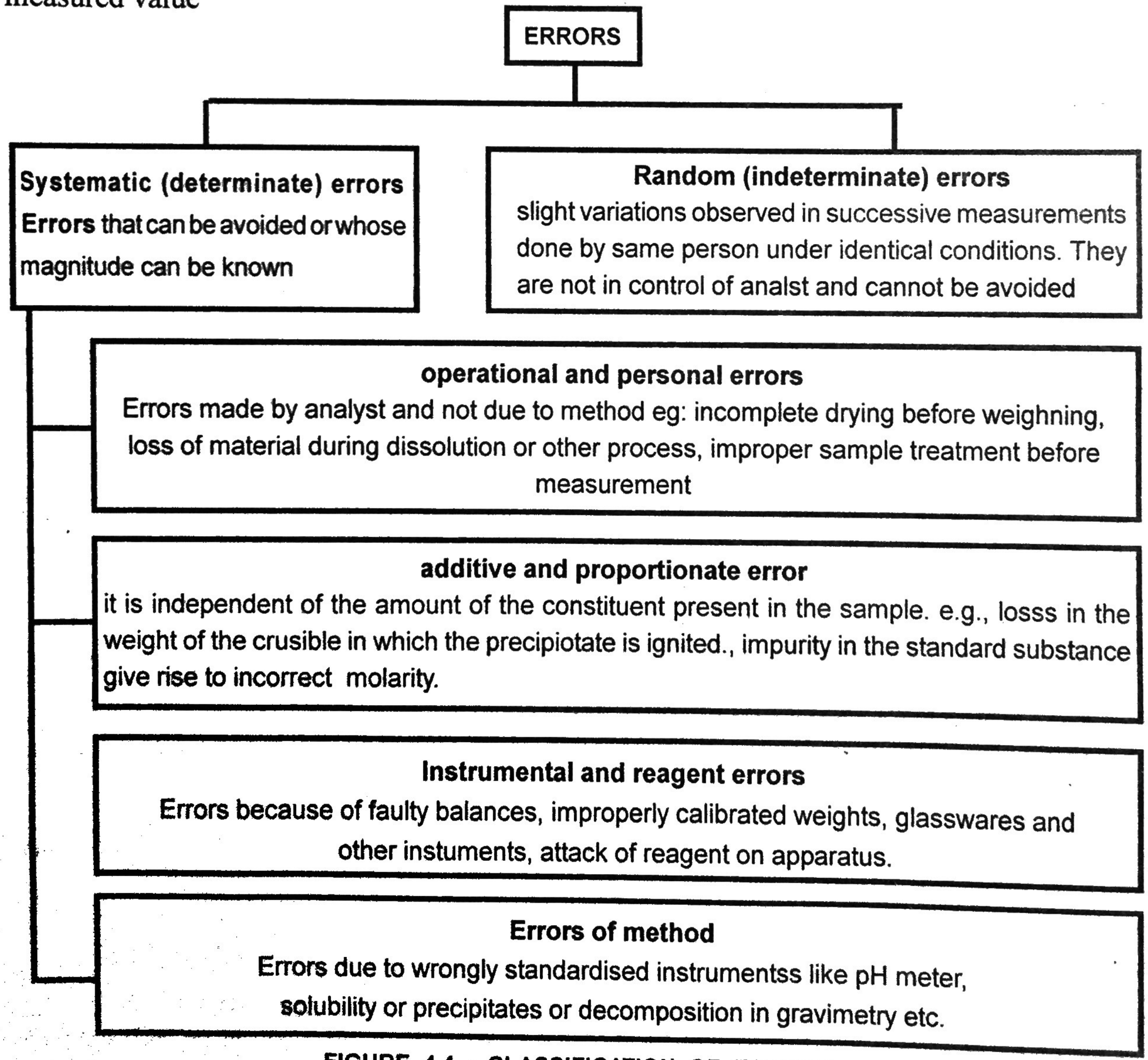
4.9 Limitations of Analytical Methods :

The function of analyst is to obtain his result as near to the true value as possible by the proper selection of the method. He must have the knowledge of accuracy and precision of the method used as well as the sources of errors. Which can be introduced every step during the analysis otherwise the confidence level with the results obtained will be very less.

Quantitative analysis requires sound knowledge of chemistry involved during testing, possibilities of interferences and statistical distribution of values. So that exact valuation of the answer and its accuracy with reliability can be judged.

4.10 Classification of Error :

Definition : "**Error** can be defined as the difference between true value and the measured value"



ERRORS

Systematic (determinate) errors
Errors that can be avoided or whose magnitude can be known

Random (indeterminate) errors
slight variations observed in successive measurements done by same person under identical conditions. They are not in control of analst and cannot be avoided

operational and personal errors
Errors made by analyst and not due to method eg: incomplete drying before weighning, loss of material during dissolution or other process, improper sample treatment before measurement

additive and proportionate error
it is independent of the amount of the constituent present in the sample. e.g., losss in the weight of the crusible in which the precipiotate is ignited., impurity in the standard substance give rise to incorrect molarity.

Instrumental and reagent errors
Errors because of faulty balances, improperly calibrated weights, glasswares and other instuments, attack of reagent on apparatus.

Errors of method
Errors due to wrongly standardised instrumentss like pH meter, solubility or precipitates or decomposition in gravimetry etc.

FIGURE 4.4 : CLASSIFICATION OF ERRORS

4.11 Accuracy and Precision :

The **accuracy** of a determination is defined as "the degree of closeness between the measured value and the true or most probable value."

Accuracy can be expressed in terms of absolute on relative error.

Systematic errors cause a constant error and hence affect the accuracy of a result.

Precision :

- The precision of a determination is defined as "the degree of closeness between a series measured value of the same property."
- Accuracy Deals With Correctness Of Measurement While Precision Deals With Reproducibility of Measurement.

4.11.1 Question :

"Precision Always Accompanies Accuracy But High Degree Of Precision Does Not Mean Accuracy"

The above statement can be explained using following example:Two analysts are analyzing a sample. The sample contains 49.10% ± 0.02% of a constituent 'A'.

The four results obtained by first analyst are as shown below :

First result	**Second result**	**Third result**	**Fourth result**	**Arithmetic mean (Average value)**	**Correct value**
49.01	49.25	49.08	49.14	49.12	49.10% ± 0.02%

The four results obtained by second analyst are as shown below :

First result	**Second result**	**Third result**	**Fourth result**	**Arithmetic mean (Average value)**	**Correct value**
49.40	49.44	49.42	49.42	49.42	49.10% ±0.02%

- **Repeatable and reproducible analysis:**

If the determinations of the values are done on the same day and identical laboratory conditions in a very short time gap, it is called **repeatable analysis.**

If the determinations of the values are done on different days and different laboratory conditions, it is called **reproducible analysis.**

4.11.2 Absolute Error (E) :

It is the difference between measured value and true value with its sign. It can be positive or negative. Mathematically,

$$E = x_i - x_t$$

x_i : experimental result (measured value)

x_t : true value

4.11.3 Relative Error (Er) :

It is the ratio of absolute error to the true value in percentage. Mathematically,

$$E_r = \frac{x_i - x_t}{x_t} \times 100$$

It also carries sign i.e. it can be positive on negative.

4.12 Minimization of Errors :

Systematic errors can be reduced by following methods:

1. **Calibration of apparatus and application of corrections :** All instruments (weights, flasks, burettes, pipettes, etc.) should be calibrated, and the appropriate corrections applied to the original measurements. If the error cannot be eliminated, one should apply a correction for the effect that it produces.

2. **Running a blank determination :** Blank determination means carrying out a separate determination without the sample under exactly the same experimental conditions as in the actual analysis of the sample. By doing this, one can find out the effect of the impurities introduced because of the reagents and vessels. A large value of blank reading increases the uncertainty in true value and reduces the precision of analysis. Hence, blank determination should have small values.

3. **Running a control determination :** Control determination means carrying out a determination on a standard substance under exactly the same experimental conditions.

The weight of the constituent in the unknown can then be calculated from the following equation :

$$\frac{\text{Result found for standard substance}}{\text{Result found for unknown sample}} = \frac{\text{Weight of constituent in standard}}{x}$$

Where, x is the weight of the constituent in the unknown.

4. **Independent methods of analysis :** The error minimization can also be done by comparing the results of the analysis carried out in two entirely different manners.

For example :

(a) Determination of iron can first be determined gravimetrically by precipitation as iron (III) hydroxide after removal the interfering elements, followed by ignition of the precipitate to iron (III) oxide. It can be also be determined titrimetrically by reduction to the iron (II) state, and titration with a standard solution of potassium dichromate or cerium (IV) sulphate.

If the results obtained by the two different methods are in close agreement, it is understood that the values are correct within small limits of error.

5. **Running parallel determinations :** In order to check the precision of results, two or three simultaneous determinations are done. If these results are in close agreement with each other, one can say that the results are in good precision. If larger variations are seen in the results, the determinations must be repeated until satisfactory agreement is obtained. However, the accuracy of the results cannot be checked by performing parallel determinations.

6. **Standard addition :** Here, a known amount of the constituent to be determined is added to the sample, which is then analyzed for the total amount of constituent present. The difference between the analytical results for samples with and without the added constituent gives the recovery of the amount of added constituent. If the recovery is satisfactory, the accuracy of the procedure is high. This method is usually applied to procedures like polarography and spectrophotometry.

7. **Internal standards :** It involves addition of a fixed amount of a reference material (the internal standard) to a series of known concentrations of the material to be measured. The ratios of the physical value (absorption or peak size) of the internal standard and the series of known concentrations are plotted against the concentration values. This should give a straight line. This procedure is of particular value in spectroscopic and chromatographic determinations.

8. **Amplification methods :** These methods are used when amount of material to be measured is very small and is beyond the limit of the apparatus. Here, this small amount of material is reacted in such a way that every molecule produces two or more molecules of some other measurable material, the resultant amplification then brings the property to be measured in the range of the measurement of the apparatus or the method.

9. **Isotopic dilution :** A known amount of the element to be determined, containing a radioactive isotope, is mixed with the sample and the element is isolated in a pure form (usually as a compound), which is then determined. The radioactivity of the isolated material is measured and compared with that of the added element. The weight of the element in the sample is then calculated.

4.13 Significant Figure :

The term 'digit' denotes any one of the numbers; 0,1,2,3....9.

A significant figure is a digit which shows the amount of the quantity in the place in which it stands.

"The number of digits which are needed to express the precision of the measurement from which the number was derived are known as significant figures."

For example, the digit zero is a significant figure except when it is the first figure in a number. Thus in the quantities **1.2680** g and **1.0062** g the zero is significant, but in the quantity 0.0025 kg the zeros are not significant figures; they serve only to locate the decimal point and can be removed by selecting proper units, i.e. 2.5 g. The first two numbers contain five significant figures, but 0.0025 contains only two significant figures.

4.13.1 Question :

Mention the correct significant figures for the given number and indicate which zeros are significant.

(a) 0.216 → three significant figure, zero is not significant figure

(b) 90.7 → three significant figure, zero is significant figure

(c) 800.0 → four significant figure, all zeros are significant

(d) 0.0670 → three significant figure, only last zero is significant

4.14 Rounding off :

If the last significant figure digit is '5' or greater than '5' (> 5) then the number is rounded up to the next higher digit.

For example :

9.47 can be round off to 9.5

9.43 can be round off to 9.4

4.15 Mean, Median, Standard Deviation :

"**Mean** is the average value of the given set of data. It is expressed as $\bar{X}$. It is also known as arithmetic mean $(\bar{X})$.

$$\bar{X} = \frac{x_1 + x_2 + x_3 + \ldots\ldots + x_n}{n} = \frac{\sum_{i=1}^{i=n} X_i}{n}$$

Median is the positional mean or central value of the given set of data. It is symbolised by M. To calculate M, arrange all data in an order i.e. either ascending or descending order. Then if n is odd number, then the middle value is the median.

But if n is even, then after arranging all data in an order. Median is the average of two middle ones.

i.e. $$M = \frac{\left(\frac{n}{2}\right)^{th} \text{data} + \left(\frac{n}{2}+1\right)^{th} \text{data}}{2}$$

- **Range :** It is the difference between largest value and smallest value of the data. It is the variability of the set of results.
- **Standard deviation / Root mean square deviation :**

It is defined as the square root of the mean of the sum of the squares of the differences between the values and the mean of those values (this is expressed mathematically below) and is of particular value in connection with the normal distribution

For '*n*' number of observations,

$$x_1, x_2, x_3, \ldots.. x_{n-1}, x_n$$

The arithmetic mean is given by

$$\overline{x} = \frac{x_1 + x_2 + \ldots. + x_{n+1} + x_n}{n}$$

The standard deviation is given by

$$s = \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots. + (x_n - \overline{x})^2}{n-1}}$$

In the equation the denominator is $(n - 1)$ rather than n when the number of values is small.

The equation may also be written as,

$$S = \sqrt{\frac{\Sigma (x_i - \overline{x})^2}{N-1}}$$

The square of the standard deviation is called the variance. $V = s^2$.

A measure of precision known as the relative standard deviation (RSD) is given by

$$RSD = \frac{S}{\overline{x}}$$

Coefficient of variation (CV) is the percentage of relative standard deviation (RSD)

$$CV = \frac{S \times 100}{\overline{x}}$$

Example 1 : Analyses of a sample of iron ore gave the following percentage values for the iron content: 7.08, 7.21, 7.12, 7.09, 7.16, 7.14, 7.07, 7.14, 7.18, and 7.11. Calculate the mean, standard deviation and coefficient of variation for the values.

Results (x_i)	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
7.08	− 0.05	0.0025
7.21	0.08	0.0064
7.12	− 0.01	0.0001
7.09	− 0.04	0.0016
7.16	0.03	0.0009
7.14	0.01	0.0001
7.07	− 0.06	0.0036
7.14	0.01	0.0001
7.18	0.05	0.0025
7.11	− 0.02	0.0004
$\Sigma x_i = 71.30$ Mean $\bar{x} = 7.13$		$\Sigma(x_i - \bar{x})^2 = 0.0182$

$$\text{Mean} = \bar{x} = \frac{\Sigma x_i}{N}$$

$$= \frac{7.08 + 7.21 + \ldots + 7.18 + 7.11}{10}$$

$$= \frac{71.30}{10}$$

$$= 7.13$$

$$\text{Standard deviation} = s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{N-1}}$$

$$s = \sqrt{\frac{0.0182}{9}}$$

$$= \sqrt{0.0020}$$

$$= \pm\ 0.045$$

$$\text{Coefficient of variation} = \text{CV} = \frac{S}{\bar{x}} \times 100$$

$$= \frac{0.045}{7.13} \times 100$$

$$= 0.63$$

Example 2 : Calculate the mean and standard deviation of the following set of analytical results : 15.67 g, 15.69 g, 16.03 g.

Solution :

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
15.67	0.13	0.0169
15.69	0.11	0.0121
16.03	0.23	0.0529
$\Sigma x_i = 47.39$	$\Sigma x_i - \bar{x} = 0.47$	$\Sigma(x_i - \bar{x})^2 = 0.0819$

$$\bar{x} = \frac{\Sigma x_i}{N} = \bar{x} = \frac{47.39}{3} = 15.80$$

$$s = \sqrt{\frac{0.0819}{3-1}} = 0.20\text{g}$$

Example 3 : The following replicate weighings were obtained; 29.8 mg, 30.2 mg, 28.6 mg and 29.7 mg. Calculate the standard deviation.

Solution :

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
29.8	0.2	0.04
30.2	0.6	0.36
28.6	1.0	1.00
29.7	0.1	0.01
Σ 118.3	Σ 1.9	Σ 1.41

$$\bar{x} = \frac{118.3}{4} = 29.6$$

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N-1}} = \sqrt{\frac{1.41}{4-1}} = \pm\ 0.69\ \text{mg}$$

$$\text{RSD} = \frac{S}{\bar{X}} = \frac{0.69}{29.6} = 0.0233\ \text{mg}$$

$$\text{CV} = \frac{S \times 100}{\bar{X}} = 2.33\ \text{mg}$$

Example-4 : The normality of a solution is determined by four separate titrations and the results are 0.2041, 0.2049, 0.2039, 0.2043. Calculate mean, median, range, average deviation, relative std. deviation, ~~std. deviation,~~ coefficient of variation.

$$\text{Mean } (\bar{X}) = \frac{\sum_{i=1}^{i=n} x_i}{N} = \frac{0.2041 + 0.2049 + 0.2039 + 0.2043}{4} = 0.2043$$

Median (M) : N = 4 So the equation,

$$M = \frac{\left(\frac{n}{2}\right)^{\text{th}} \text{data} + \left(\frac{n}{2}+1\right)^{\text{th}} \text{data}}{2}$$

$$= \frac{0.2041 + 0.2043}{2}$$

$$= 0.2042$$

$$\text{Range (R)} = \text{Highest value} - \text{Lowest value}$$
$$= 0.2049 - 0.2039$$
$$= 0.0010$$

$$\text{Std. deviation} = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{N-1}} = 0.0004$$

$$\text{Coefficient of variation (C.V.)} = \frac{S}{\bar{X}} \times 100$$
$$= \frac{0.0004}{0.2043} \times 100$$
$$= 0.2\%$$

4.16 Distribution of Random Errors :

Random errors can not be avoided or minimized but can be monitored and its magnitude can be measured. These accidental errors will follow a random distribution and therefore mathematical laws of probability can be applied to achieve some conclusion.

Indeterminate errors (random errors) follow a Normal Distribution Curve or Gaussian Distribution Curve, as shown in the figure.

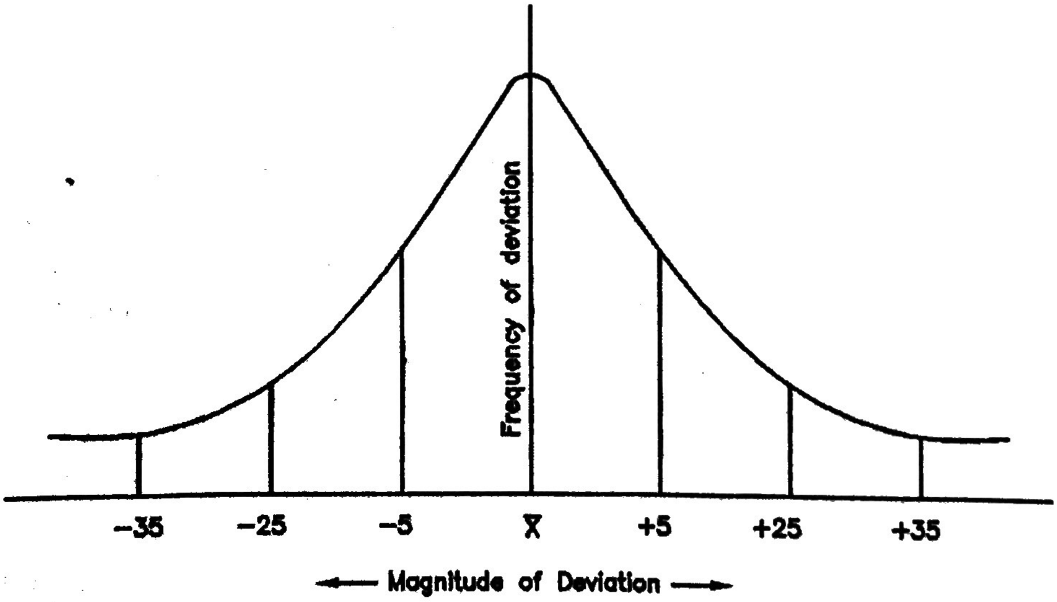


FIGURE 4.5 : DISTRIBUTION OF RANDOM ERRORS

From the curve it is clear that almost

68% of results fall between the range $\bar{X} \pm S$ S,

95% of results fall between $\bar{X} \pm 2S$,

99% and more results fall between $\bar{X} \pm 3S$

Where 'S' indicates standard deviation and $\bar{X}$ represents its mean value. This curve is symmetrical on both the sides.

4.17 Reliability of Results : (Q-test) :

Analysis of the results can be considered in two main categories : (a) the reliability of the results; and (b) comparison of the results with the true value or with other sets of data.

A most important consideration is to be able to arrive at a sensible decision as to whether certain results may be rejected. It must be concerned that values should be rejected only when a suitable statistical test has been applied or when there is an obvious chemical or instrumental reason that could justify exclusion of a result.

Q-test the parameter Q is called the rejection quotient and is defined as;

$$Q = \frac{[\text{Questionable Value} - \text{Nearest Value}]}{\text{Largest Value} - \text{Smallest Value}}$$

For the calculations; all the results are to be arranged in an order and then the nearest value can be considered. The questionable data is rejected if the calculated value of Q_{cal} is higher than the critical value of Q (obtained from the Q-table) for the given number of observations. i.e. if $Q_{cal} > Q_{crit}$, then questionable value is rejected.

Example 5 : The following values were obtained for the determination of cadmium in a sample of dust : 4.3, 4.1, 4.0, 3.2 $\mu g\ g^{-1}$ Should the last value, 3.2, be rejected?

The Q test may be applied to solve this problem. **Qcrit = 0.831**

$$Q = \frac{[\text{Questionable Value} - \text{Nearest Value}]}{\text{Largest Value} - \text{Smallest Value}}$$

$$Q = \frac{[3.2 - 4.0]}{4.3 - 3.2} = \frac{0.8}{1.1} = 0.727 \quad \text{Here, Qcal} < \text{Qcrit}$$

If the calculated value of Q exceeds the critical value, then the questionable value may be rejected. In this example Q calculated is 0.727 and Q critical, for a sample size of four, is 0.831. Hence the result 3.2 $\mu g\ g^{-1}$ should not be rejected.

Example 6 : If, however, in the above example, three additional measurements were made, with the results: 4.3, 4.1, 4.0, 3.2, 4.2, 3.9, 4.0 $\mu g\ g^{-1}$. Qcrit = 0.570

$$Q = \frac{[3.2 - \overset{3.9}{\cancel{4.0}}]}{[4.3 - 3.2]} = \frac{0.7}{1.1} = 0.636 \quad \text{Here,} \quad Qcal > Qcrit$$

The value of Q critical for a sample size of seven is 0.570, so in this case the value 3.2 $\mu g\ g^{-1}$ is rejected.

4.18 Confidence Interval (Limit) :

When a small numbers of observations is made, then standard deviation of these results can not give a measure of how close $\bar{X}$ (mean) is to the true mean value (μ). But it is possible to calculate a confidence interval to estimate the range within which the true mean can be found. This limit is known as confidence interval and can be calculated as;

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{n}}$$

μ : true mean value

$\bar{X}$: average (mean) value for given set of data

S : standard deviation

n : number of data

t : parameter value obtained from t-table at given degrees of freedom (n - 1)

Example 7 : The mean $(\bar{X})$ for four determination of alloy analys is was 8.27%, Std. deviation was S = 0.17% calculate confidence interval at 95% confidence level. t = 3.18 for (n − 1) = 3 degrees of freedom.

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{n}}$$

$$= 8.27 \pm \frac{(3.18)\,(0.17)}{\sqrt{4}}$$

$$= 8.27 \pm 0.27\%$$

So, the confidence interval is 8.00 to 8.54%

Note : if we increase the value of N = 12 then t = 2.20 and confidence interval is

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{N}}$$

$$= 8.27 \pm \frac{(2.20)\ (0.17)}{\sqrt{12}}$$

$$= 8.27 + 0.11\%$$

So, the confidence interval is 8.16 to 8.38%. Thus by increase 'N', we can decrease the range and can increase the probability of finding true mean.

Example 8 : The amount of mercury was determined in the given samples as follows :

1.80, 1.58, 1.64 ppm Hg

Its $\bar{X}$ = 1.673 and S = 0.0258. Calculate confidence interval at 95% confidence level with t = 1.96 for given set of data. Conclude your answer for N = 10.

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{N}}$$

$$= 1.673 \pm \frac{(1.96)\ (0.0258)}{\sqrt{3}}$$ ✓

$$= 1.673 \pm 0.029$$

So confidence interval is 1.644 to 1.702.

$$\mu = \bar{X} \pm \frac{ts}{\sqrt{N}}$$

$$= 1.673 \pm \frac{1.96\ (0.0258)}{\sqrt{10}}$$

$$= 1.673 + 0.016$$

So, confidence interval is 1.657 to 1.689

"By increasing number of replicate analysis, confidence interval decreases and we get more probability to find the true value."

Note : (1) If 'N' increases μ decreases. So it is advised to increase the replicate analysis during experiments, which will lower the range of finding the true value.

(2) If 'S' increase, the μ increases. So always when deviation is more, the range to find out true value will be broad. Thus, always try to keep std. deviation as small as possible.

4.19 Comparison of Results :

Comparison of the values obtained from a set of results can be done by two ways :

(a) either with the true value or (t - test)

(b) with the other set of data to explain whether the process is precise or not; superior or not. (F - test).

These test methods requires the information of degrees of freedom. In statistics, this is the number of independent values necessary to determine the statistical quantity.

Thus, a sample of 'N' values will have 'N' degrees of freedom. Whereas the sum $\sum(x_i - \bar{X})^2$ is considered to have N-1 degrees of freedom. Because for any defined value of $\bar{X}$ only N-1 values can be freely assigned, the N^{th} being automatically defined from the other values.

4.19.1 Student's t-test :

It is used for small samples (with limited quantity of data / results). Its purpose is to compare the mean of sample with some standard value and to express some level of confidence for comparison. The equation for t-test is;

$$t = \frac{|\bar{X} - \mu| \sqrt{N}}{S}$$

$\bar{X}$: Mean value

μ : true value (standard value)

N = Number of data

S = Standard deviation

This equation gives t value termed as 'tcal' means calculated t value. Then it is compared with t tab (given with numerical) which is being taken from t-table (appendix) for the given degrees of freedom and at given confidence level.

If, $t_{cal} > T_{tab}$ then, the results are not significant.

If, $t_{cal} < t_{tab}$ then, the results are significant or reliable

Example 8 : If $\bar{X}$ of 12 determinations is 8.37, true mean value is $\mu = 7.91$ and $S = 0.17$, say whether the results are significant or not with the three mean value. Given $t = 2.20$ at 95% confidence level for 11 degrees of freedom.

$$t = \frac{|\bar{X}-\mu|\sqrt{N}}{S}$$

$$= \frac{|8.37-7.91|\sqrt{12}}{0.17}$$

$$= 9.4$$

$t_{cal} = 9.4$

$t_{tab} = 2.20$

So, $t_{cal} > t_{tab}$. Thus the results are not significant with the true mean value.

4.19.2 F-test :

It is used to compare the precision of two sets of data (results). It can be results of two methods for the same samples, results of two analyst, two laboratories etc. .... The comparison two mean and standard deviation gives comparison of precision. The equation for F-test is;

$$F = \frac{S_A^2}{S_B^2}$$

Where S_A : Standard deviation of method-A with higher value.

S_B : Standard deviation of method-B with lower value.

Note : For the calculation of F-test, always $S_A > S_B$, So that F_{cal} is always greater than one.

After calculation F_{cal} is compared with F_{tab} for the given degrees of freedom, derived from F_{table} (Appendix).

If $F_{cal} > F_{tab}$, the two sets of results are not significant consistent or comparable, and

If $F_{cal} < F_{tab}$ then the sets of results or mean are significant.

Example : $S_A = 0.210$ for 11 determination and $S_B = 0.641$ for 13 determinations was derived. From F-test, calculate if there is any significant difference found from the two sets of data.

$F_{tab} = 2.91$ for 95% confidence level.

$$F = \frac{S_A^2}{S_B^2} \quad \ldots\ldots \text{ Where } S_A > S_B$$

$$= \frac{(0.641)^2}{(0.210)^2} = \frac{0.411}{0.044} = 9.4 \qquad F_{cal} > F_{tab}$$

$F_{cal} = 9.4$ which is greater than F_{tab}, So the two sets of data are different and not comparable.

Example : Compare the newly developed method with standard method from the given results; for % Nickel in the sample. $F_{tab} = 5.19$ for given degrees of freedom.

	New Method	**Standard Method**
Mean ($\bar{X}$)	7.85	8.03
Standard Deviation (S)	± 0.130	+ 0.095
Number of samples (N)	5	6

$$F = \frac{S_A^2}{S_B^2} \quad \ldots\ldots\ldots \; S_A > S_B$$

$$= \frac{(0.130)^2}{(0.095)^2}$$

$$= 1.87$$

$\therefore F_{cal} = 1.87$

Now $F_{tab} = 5.19$ which is given to us

$\therefore F_{cal} < F_{tab}$ and hence the new method is significant with standard method or comparable with standard method.

Question Bank

- **Multiple Choice Question :**

1. The analysis in which we finds out selected constituents of the sample is known as.
   (a) Proximate analysis (b) Partial analysis
   (c) Trace constituent analysis (d) Complete analysis
2. The analysis in which the amount of sample taken is less than 0.00001gm is known as
   (a) Sub micro analysis (b) Partial analysis
   (c) Ultra micro analysis (d) Micro trace analysis
3. The analysis in which the amount of sample taken is 0.1gm to 0.0001gm is known as
   (a) Sub micro analysis (b) Micro analysis
   (c) Ultra micro analysis (d) Micro trace analysis
4. The instrumental technique in which the measurement of difference in temperature between sample and reference material is done is called
   (a) TGA (b) DTA (c) DSC (d) Electrogravimetry
5. The method based on measurement of potential of an electrode in equilibrium with an ion to be studied is known as
   (a) Voltametry (b) Coulometry (c) Potentiometry (d) Chromatography
6. Which one of following is method based on study of movement of a substance in a medium under controlled conditions
   (a) Voltametry (b) Mass spectroscopy
   (c) Infra red spectroscopy (d) visible spectroscopy
7. Errors made by analyst and not due to method is called as
   (a) random error (b) operational error
   (c) instrumental error (d) Errors of method
8. How many significant figure is present in the data: 0.1025 ?
   (a) 1 (b) 2 (c) 3 (d) 4
9. The difference between the observed value and the true value is called.
   (a) Reproducibility (b) Precision
   (c) Accuracy (d) Error
10. If the calculated value of Q exceeds the critical value, then the questionable value may
    (a) Not be rejected (b) Be rejected
    (c) Both of this (d) None of this.

11. When the amount of all the constituents is determined, then the method is
(a) Proximate analysis
(b) Partial analysis
(c) Trace constituent analysis
[✓ (d) Complete analysis]

12. Volumetric analysis and Gravimetric analysis are the types of :
(a) thermal analysis
[✓ (b) quantitative analysis]
(c) optical method
(d) all of these

13. The amount (quantity) of the substance taken for the macro analysis is
(a) Sub micro analysis
[✓ (b) Macro analysis]
(c) Ultra micro analysis
(d) Micro trace analysis

14. If the error arises due to voltage fluctuation then the type of error is
[✓ (a) systematic error]
(b) random error
(c) determinate error
(d) none of these

15. Gravimetry and volumetry are ....
[✓ (a) Chemical analysis]
(b) Instrumental analysis
(c) Both of these
(d) None of these

16. For macro analysis what amount of sample is required ?
[✓ (a) 0.1 gm and more]
(b) 0.1 gm to 0.001 gm
(c) 0.001 gm to 0.0001 gm
(d) less than 0.0001 gm

17. The process of identifying the components of materials and the impurities present in the sample is called ....
(a) Spectroscopic method
[✓ (b) Qualitative method]
(c) Quantitative method
(d) IR method

18. Which of the following is based on spectroscopic properties measurements ?
(a) Chromatography
(b) Potentiometry
[✓ (c) Flame photometry]
(d) Titration

19. Which of the following method is based on amount of sample ?
(a) Proximate Analysis
(b) Partial Analysis
(c) Complete Analysis
[✓ (d) Macro Analysis]

20. Which of the following method involves electrical measurements ?
(a) Retractometry
(b) Absorption method
[✓ (c) Voltametry]
(d) Chromatography

21. The degree of closeness between measured value and true value is ....
[✓ (a) Accuracy]
(b) Precision
(c) Method
(d) Analysis

22. The quantity of the sample taken in sub-micro analysis is ....
(a) 10^{-4} gm or less
(b) 10^{-3} gm to 10^{-12} gm
(c) 10^{-2} gm to 10^{-1} gm
[✓ (d) 10^{-3} gm to 10^{-4} gm]

23. The sample taken for the analysis should always be ______ .
(a) impure [✓ (b) homogeneous] (c) turbid (d) liquid

24. The precision of the result is also known as ______ .
(a) Accuracy (b) Error
[✓ (c) Reproducibility of result] (d) All of these

25. The quantity for nano-trace analysis is ______ .
[✓ (a) 10^{-7} to 10^{-10} ppm] (b) 0.001 gm
(c) 1.00 gm (d) 10.0 gm

26. The positional mean for the given set of data is ______
(a) Arithmatic Mean [✓ (b) Median]
(c) Standard Deviation (d) Relative Error

27. Random Errors are spreaded over both the sides of mean value and follows ______ pattern.
[✓ (a) Gaussian Distribution curve] (b) Straight line
(c) both of these (d) none of these

28. If we increase number of replicate analysis, the range of finding true value ...
(a) increases [✓ (b) decreases]
(c) remains constant (d) decreases linearly

29. Which of the following test method/s carry out comparison of mean value with true result ?
(a) F-test [✓ (b) t-test] (c) Q-test (d) All of these

30. Which of the following method does comparison of mean of two sets of Data ?
[✓ (a) F-test] (b) t-test (c) Q-test (d) All of these

31. For calculations of t-test & F-test, degrees of freedom can be calculated from which of the following equation ?
(a) $N + 1$ (b) $N - 1$ (c) $N^2 + 1$ (d) $(N + 1)^2$

32. Which one of the following represents the most correct answer of central value ?
(a) Arithmatic mean (b) Median
(c) Variance (d) Range

33. Which of the following data has three significant figure ?
(a) 1.379 (b) 1.37 (c) 1.3790 (d) 1.37900

34. Number of significant figure for the data 0.00250 is ______ .

(a) 3 (b) 5 (c) 2 (d) 6

35. Rounding off the data 1.3798 to three significant figures gives ______ .

(a) 1.38 (b) 1.37 (c) 1.39 (d) 1.40

- **Short Questions : (2 marks each)**

1. Define : Quantitative Analysis and Qualitative Analysis, Accuracy and Precision, repeatable and reproducible analysis, absolute error and relative error.
2. List out factors affecting selection of analytical method for the analysis.
3. Give classification of chemical analysis based on data to be generated.
4. "Precision always accompanies accuracy but high degree of precision does not mean accuracy" justify the statement.
5. Define sampling and real sample.
6. Write the number of significant figures against the data :
   1.2680, 1.0062, 0.0025, 2.5, 1.0001000, 0.000009, 1.23456, 7.77007700, 1.002, 0.02370, 328.4
7. Discuss advantages of chemical method.
8. Discuss limitations of chemical method.
9. Discuss advantages of instrumental method.
10. Discuss limitations of instrumental method.
11. Explain the terms :
    (a) Systematic error (Determinate error) & Random error (Indeterminate error)
    (b) Accuracy and precision.
    (c) Absolute error and relative error.
12. Discuss limitations of Analytical methods.
13. Discuss mean of the given set of data.
14. Explain median (positional mean) of the given set of data.
15. What is confidence interval ?
16. Discuss students t-test with its equation.
17. Discuss F-test with its equation.
18. Name three types of systematic error.
19 Name any five methods for minimizing systematic error.
20. State any five techniques (methods) for removal of interferences.

- **Explain the following :** (5 Marks)

1. Define the term chemical analysis and discuss applications of chemical analysis.
2. Discuss on : Stages of Analysis.
3. Give complete classification of chemical analysis based on data to be generated.
4. List out different methods for the minimization of systematic error and explain on any three of them.
5. Define : Error. Give complete classification of error.
6. Give classification of chemical analysis based on amount of sample.
7. Discuss advantages & limitations of chemical and instrumental methods of analysis.
8. Explain following methods for minimization of systematic error.
   (a) Running Blank Determination
   (b) Independent Methods of analysis
   (c) Amplification method
   (d) Internal standard method
   (e) Standard addition methods
9. What is sampling ? Discuss sampling of solid, liquid and gas samples.
10. Discuss on interferences and any four methods to remove it from the sample.
11. Discuss comparison of results with required equations.

- **Numericals for practice :**

1. During iron estimation of the sample following results are obtained.
   8.0, 8.2, 8.4, 8.5, 9.6 ppm. Check wether the questionable value 9.6 should be rejected or not Q_{crit} = 0.710.
2. Calculate the mean, standard deviation variance, coefficient of variation relative standard deviation for the following results.
   86.1, 85.2, 85.8, 85.7, 85.5, 85.3.
3. Calculate mean, median and standard deviation for the given set of data.
   0.256, 0.268, 0.252, 0.262.
4. The following results are obtained for % Ni in an alloy.
   7.72, 7.86, 7.54, 7.58, 7.62, 7.66, 7.05.
   Check the reliability of que$\overset{\text{st}}{\text{r}}$ionable value 7.05 with $Q\overset{\text{crit}}{\text{uit}}$ = 0.680 for given at.
5. Assign number of significant figures for the following data :
   0.05900 → 4
   0.00003 → 1

39400 → 5
70800.00 → 7
7.080×10^4 → 4
0.0020810 → 5
15860 → 5
5.08×10^4 → 3
123456 → 6

6. From the following results, calculate mean, standard deviation, relative standard deviation and variance.

   29.8, 30.2, 28.6, 29.7%

   [Hint : $\bar{x}$ = 29.6%, S = 0.69%, RSD = 1.1%]

7. The tin and zinc contents of a brass sample are analyzed with the following results :

   33.27, 33.37, 33.34%

   Calculate mean, standard deviation, relative standard deviation and the coefficient of variance, for the data.

8. Replicate samples of a silver alloy are analyzed and determined to contain 95.67, 95.61, 95.71, 95.60%. Calculate mean, standard deviation, variance, relative standard deviation, coefficient of variance.

9. The following results are obtained in a particular concentration solution. From Q-test, conclude if any of the result is required to be discarded, at 95% confidence limit.

   0.1067, 0.1071, 0.1066, 0.1050, Q tab = 0.829

10. The following are results obtained from alloy analysis to get %Fe.

    22.23, 22.18, 22.25, 22.09, 22.17%.

    From Q-test check whether the result 22.09 is reliable at 95% confidence limit or not ? Q tab = 0.710

11. The percentage of silica in a sample of mineral was 3.685, 3.689, 3.681, 3.692. Calculate mean, standard deviation, variance, coefficient of variance, relative standard deviation for the given data.

12. The normality of a solution is determined by four separate titrations, the results are : 0.2041, 0.2049, 0.2039, 0.2043. From these results, calculate mean, standard deviation relative standard deviation and variance.

    [$\bar{x}$ = 0.2043, S = ±0.0004, v = 0.2%]

13. Five determinations of vitamin-C content of a soft drink sample gave following results 0.218, 0.219, 0.230, 0.215, 0.220 mg/lit. Apply Q-test to see if 0.230 value can be discarded or not. Qtab = 0.64.

14. Apply Q-test for the following set of results to determine whether the outlying result should be considered or not for 95% confidence level

    (a) 41.27, 41.61, 41.84, 41.70

    (b) 7.295, 7.284, 7.388, 7.292

    (c) 85.10, 84.62, 84.70

    (d) 85.10, 84.62, 84.65, 84.70

    Note : Quit can be obtained from Q-table given in the Appendix.

15. A chemist obtained following results for his samples :

    0.084, 0.089, 0.079%

    Calculate confidence interval for given set of data when t = 4.30 at 95% confidence level for two degrees of freedom.[Hint : S = 0.005%]

16. An analyst reported the following % FeO in the given sample. For the given set of data, Calculate mean, median, standard deviation, variance, coefficient of variation.

    16.65, 16.70, 16.68, 16.60, 16.58, 16.63

17. Following samples were analysed by two methods with the help of F-test cheek any significant difference between the method is found.

    Method-A : 10.9, 10.1, 10.6, 11.2, 9.7, 10.0

    Method-B : 9.2, 10.5, 9.7, 11.5, 11.6, 9.3, 10.1, 11.2

    $\overline{X}_A = 10.4$ and $\overline{X}_B = 10.4$

    $F_{tab} = 4.95$

18. The experimental result for the given analysis are given. If the true value is 99.95%, check whether these results are acceptable at 95% confidence limit with t-test method.

    Result : 99.93, 99.87, 99.91, 99.86%

19. A new method was developed for the analysis of Fe (III) and the results are :

    0.1017, 0.1019, 0.1016, 0.1015.

    The standard result for Fe (III) is 0.1012. Calculate with t-test and check wether the new method is reliable or not ?

20. For the given data check whether outlying result is considered or not ?

    22.23, 22.18, 22.25, 22.09, 22.17

    Apply Q-test to check the answer

★ ★ ★