Unit - I

LABORATORY HYGIENE AND SAFETY

Safety measures

Laboratory is a place where a student practises what he learns in the class room. A chemistry laboratory is a place where a student has to store and handle chemicals. Chemicals are, by and large, dangerous substances. There are several chemicals which are corrosive, flammable, explosive, toxic, carcinogenic and poisonous. So a student's health will be spoiled if he or she does not know how to store and handle these substances carefully.

Hygiene means conditions or practices conductive to maintaining health. so if a student does not want to spoil his or her health while doing experiments he or she has to follow certain conditions and practices in the laboratory. Thus laboratory hygiene and safety assume importance in the life of a student who practises chemistry.

Storage and Handling of Chemicals

1. Corrosive Chemicals

Chemicals which corrode or destory gradually, skin, wood, cloth, metal etc. are called corrosive chemicals.

Examples :

Acids, alkalis, chlorine, bromine, phenols etc.

Storage :

They must be stored in corrosive resistant chambers or in pits containing sand.

Handling :

Acids should not be poured directly from their containers. Instead, a siphon or funnel must be used to transfer them from their containers without spilling them on the floor, table or person. Rubber gloves and rubberised apron may be used to avoid accidental spillage on the body or the cloth. Sodium hydroxide pellets or sticks should not be hand picked. A pair of forceps is to be used for this purpose. Chlorine and bromine must be handled infume cupboards only without allowing their vapours to come into contact with nose, eyes or skin.

2. Flammable chemicals

Chemicals which catch fire on exposure to air or on heating over a naked flame or in a warm surrounding are called flammable chemicals.

Examples :

Ether, benzene, acetone, carbon disulphide, alcohol, sodium, potassium, etc.,

Storage :

They should be kept in bottles with tight lids in a cool place. Sodium and potassium are to be stored under kerosene only.

Handling :

They should not be handled near burning burners, heaters etc. They should not be heated directly, over a flame. They should be heated in a flask using a reflux condenser, on a water bath or steam bath. While heating them in this way the near by burners must be put off. Buring Sodium or Potassium metal can be extinguished using foam or sand. Water should not be used.

3. Explosive chemicals

Chemicals which explode violently on heatng, grinding or pressing are called explosive chemicals.

Examples :

Chlorates, perchlorates, nitrates, ethers, peroxides, poly nitro compounds etc.

Storage :

They should be kept in such a way that there is no empty space above them in the container in which they are kept. They must be stored in a cool place.

Handling :

They must be handled with all windows open and wit exhaust fan on. They must be very carefully heated or ground.

4. Toxic chemicals.

They are substances related to poison. When taken in, they produce ill effects creating health problems.
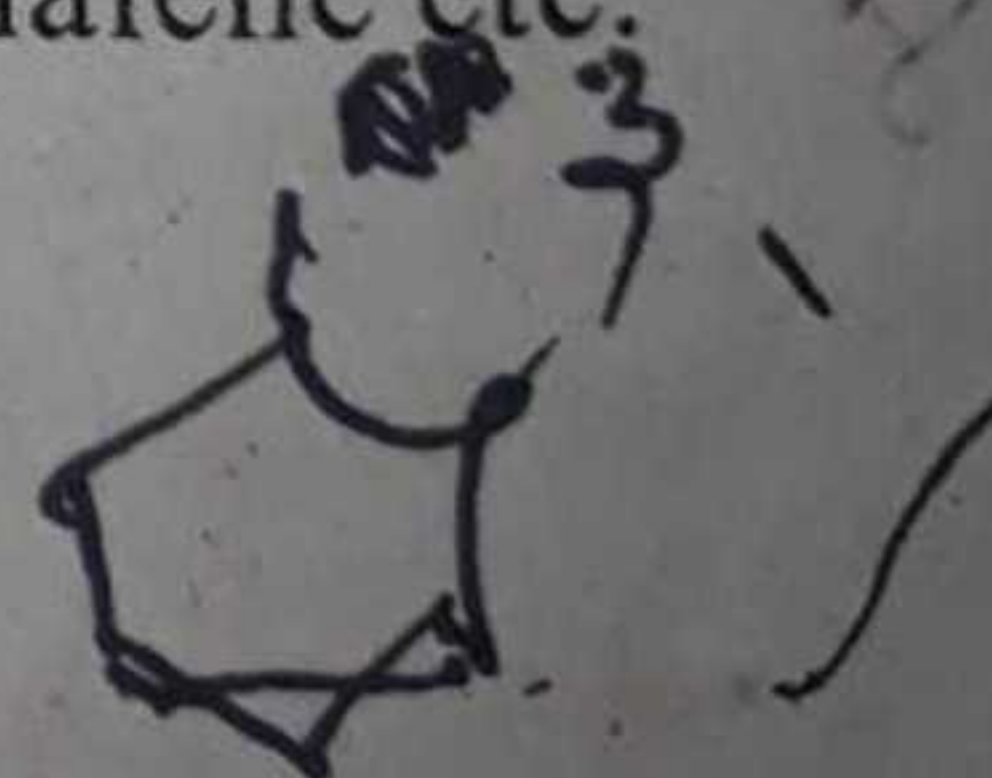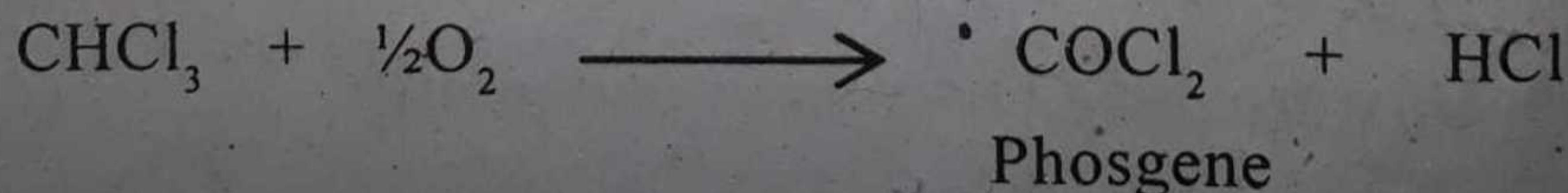
Examples :

Benzene, toluene, chloroform, carbon tetra chloride, naphthalene etc.

[Note : Almost all the chemicals are toxic].

Storage :

tight lid.

They must be stored in well sealed bottles.

lid   CCl₄

Chloroform is stored in dark brown / blue bottles to prevent the formation of highly poisonous phosgene.

$$CHCl_3 + \tfrac{1}{2}O_2 \longrightarrow COCl_2 + HCl$$

Phosgene

Handling :

They must be handled with all windows open and in fume cup - boards. They must not be inhaled directly. Smellingtest must be done by keeping the test tube containing the toxic subtance away from the nose and waving the vapours towards the nose and inhaling slowly and in small quantities. They should not be handled by naked hands. Gloves must be used while handling them.

5. Carcinogenic Chemicals

Chemicals which cause cancer are called carcinogenic.

Examples :

Naphthyl amines, salts of napthyl amines, aziridine, methyl iodide, benzene (causes blood cancer), benzphyrene, thiourea, diazomethane, dimethyl sulphate, N - nitrosocompounds, etc.

Storage :

They should be stored in locked containers inside fume cup - boards. Warning lables showing a skill and a pair of crossed bones should be pasted onthe container.

Handling :

They must be handled only with gloves on. The vapours should not be inhaled. They should not be allowed to come into contact with the skin. They must be handled in fume cup - boards with the exhaust fan on.

6. Poisonous chemicals

Chemicals which when introduced into or absorbed by a living organism causes death or injury expeciclly one that kills by rapid action even in small quantity are called poisonous chemicals.

Actually their is no great difference between toxic and poisonous chemicals. We can distinguish between them. by saying that those chemicals whcih cause death realtively quickly are poisonous. Chemicals which are injurious to health in someway, which may lead to death in course of time are toxic.

Examples :

Benzene, toluene, xylenes, napthalene, anilines, $CHCl_3$, CCl_4, DDT, hydrazine, salts of Pb, selinium compounds, cyanides, tellurium compounds, chromium compounds, vanadium compounds, arsenic compounds etc.

Storage :

They must be stored in well sealed tubes. They should be labelled as poisons.

Handling :

They must be handled only with gloves on. The vapours should not be inhaled. They should not be allowed to come into contact with the skin.

University Questions

1. Give two examples corrosive chemicals.

2. Explain the methods used of the storage and handling of corrosive reagents.

3. Mention an example for a flammable chemical.

4. How are flammable chemicals stored?

5. How will you handle a flamable chemicals.

6. How is wether handled before proceeding with experiment?

7. Give a brief account on the usage and storage of flammable.

Unit - II

DATA ANALYSIS

In analytical chemistry were are going to get a lot of data. How to analyse and interpret these data? This we shall see in this chapter.

1. ERRORS IN CHEMICAL ANALYSIS

Definition of Error :

The difference between the measured value of a property and its accurate value is called the error.

Explanation :

During chemical analysis we measure the value of a particular property. E.g., we measure the weight of an object or the volume of a solution. Accurate results will be got when persons with great skill do the measurements with best instruments. This is nearly impossible. Usually the measured value of the property will never be the accurate value of the property. The difference between these two is called the error. Such errors in measurement will affect the accuracy and precision of the property that is measured. So, the analytical data so obtained becomes unreliable. In the following pages we shall study about such errors in chemical analysis.

Classification of Errors :

The errors that arise in a chemical analysis are classified into two types. They are

    i. Determinate errors

    ii. Indeterminate errors (or) Random errors.

It should be remembered that it is difficult or impossible to be certain of the type to which a given error belongs. Never the less the classification is useful for discussing analytical errors.

Determinate Errors (or) Systematic errors

These are errors which have a definite value and an assignable cause. The analyst can measure and account for these errors. These can be avoided. They are unidirectional i.e., the errors will be either more or less than the accurate value. From this, they can be identified.

Sources of these errors :

    i.   Defective instruments
    ii. Careless operation
    iii. Procedural defects

Classification :

    i.   Instruments errors.
    ii. Method errors .
    iii. Personal errors.

i. Instrument Errors :

When we use balances, weights, pipettes, burettes etc., we must make sure that they are not defective. For example, a weight marked 10g may not be 10g after all. So to avoid these errors one must use best instruments. Periodic calibration of apparatus and weights is a must. These errors may be identified by changing the instrument, the error will also change.

ii. Methods Errors :

These are introduced by defective expermental procedures. E.g.,

    a. Co - precipitation or post precipitation in gravimetric analysis.
    b. Usage of improper indicators in volumetric analysis.

These are difficult to identify. So these are the most serious of the three types of determinate errors. So to avoid these, one must be thorough with the theoretical part of the experiment.

iii. Personal Errors (or) Operative Errors

These are introduced by personal defects or carelessness. The Sources of this error are human defects in eyes, mind etc., In colorimetric

experiments errors will be introduced by a person who is colourblind. A person with defective eyes will invariably note readings erraneously. Carelessness, fatigue and improper instructions from the teacher also introduce these errors. Several mistakes may creap in E.g., Wrong calculations, wrong placement of decimals, noting wrong signs, cooking results etc. These can be avoided if one works scrupulously in the laboratory. These are identified by the fact that these errors change when the measurement is repeated.

Determinate errors may also be classified as being either

1. Constant Errors and 2. Proportional Errors.

1. Constant Errors :

These are errors whose magnitude is independent of the size of the sample taken for analysis. For example, let us say that 0.5 mg of percipitate is lost when washed with 200ml of the wash liquid. Now, if we wash 500mg of precipitate with 200 ml of wash liquid 0.5 mg of the precipitate will be lost. So the loss is $(0.5 \times 100) \div 500 = 0.1\%$. Let us assume that we wash 50mg of the precipitate with 200ml of wash liquid, here also 0.5mg will be lost. So the loss is $(0.5 \times 100)\ 50 = 1\%$.

Thus we find that a constant error will become more serious as the size of the quantity measured decreases. So to minimize the effect of constant error we have to use a large, sample. [Please remember that in our experiments in gravimetric analysis the solution to be estimate is so prepared that the weight of the precipitate is around 0.2g].

2. Proportional Errors :

These are errors whose magnitude increases or decreases in proportion to the size of the sampel taken for analysis. Invariably impuritites in the sample. If not removed, will cause a proportional error.

Correction of Determinate Errors :

Determinate instrumental errors are corrected by calibrating the

instruments concerned. The equipments are to be calibrated periodically. This is because instrumental errors arise due to wear corrosion or mistreatment.

Determinate personal errors can be minimised by care and self discipline. Instrument readings, note book entries and calculations should be checked systematically.

Determinate method errors are particularly difficult to detect. They may be corrected by one or more of the following procedure.

1. **Analysis of standard samples :**

A method may be tested for determinate error by analysis of a synthetic sample whose over all composition is known and which closely resembles the material to be tested by the particular method.

2. **Independent analysis :**

When samples to be analysed are not available in a pure state this method is used. The sample is analysed in a particular method. Then it is anlysed by a different method of established reliability.

3. **Blank determinations :**

Constant errors affecting physical measurements can be frequently evaluated with a blank determination in which all steps of the analysis are performed in the absence of the sample. The result is then applied as a correction to the actual measurement. This method is useful to correct errors that the due to the introduction of interfering contaminants from reagents and vessels employed in the analysis. This method is useful to correct titration data in volumetric analysis.

4. **By taking large sample size :**

We know that a constant error decreases as the size of the sample is increased. So, to correct such type of errors large sample size is used for analysis.

Random Errors or Indeterminate Errors or Accidental Errors

They are errors arising from uncertainities in a measurement that are

unknown and not controlled by the person doing an experiment.

Sources :

    i.   Instrument uncertainities

    ii. Method uncertainities

    iii. Personal uncertainities.

Identification :

    Indentification of indeterminate errors is difficult. Scatter of data about the mean is the effect of an indeterminate error.

Error Analysis :

    When indeterminate error or deviation from mean (dm) is ploted against its frequency (f) we get a curve as shown in figure. This bell shaped curve is called Gaussian of normal error curve. The properties of this normal curves are

i.   The frequency is maximum where the indeterminate error is nil.

ii.   There is a symmetry about this maximum, suggesting that positive and negative errors occur with equal frequency.

iii. As the magnitude of the error increases, the frequency decreases exponentially.

      - 0.4   dm   +0.4
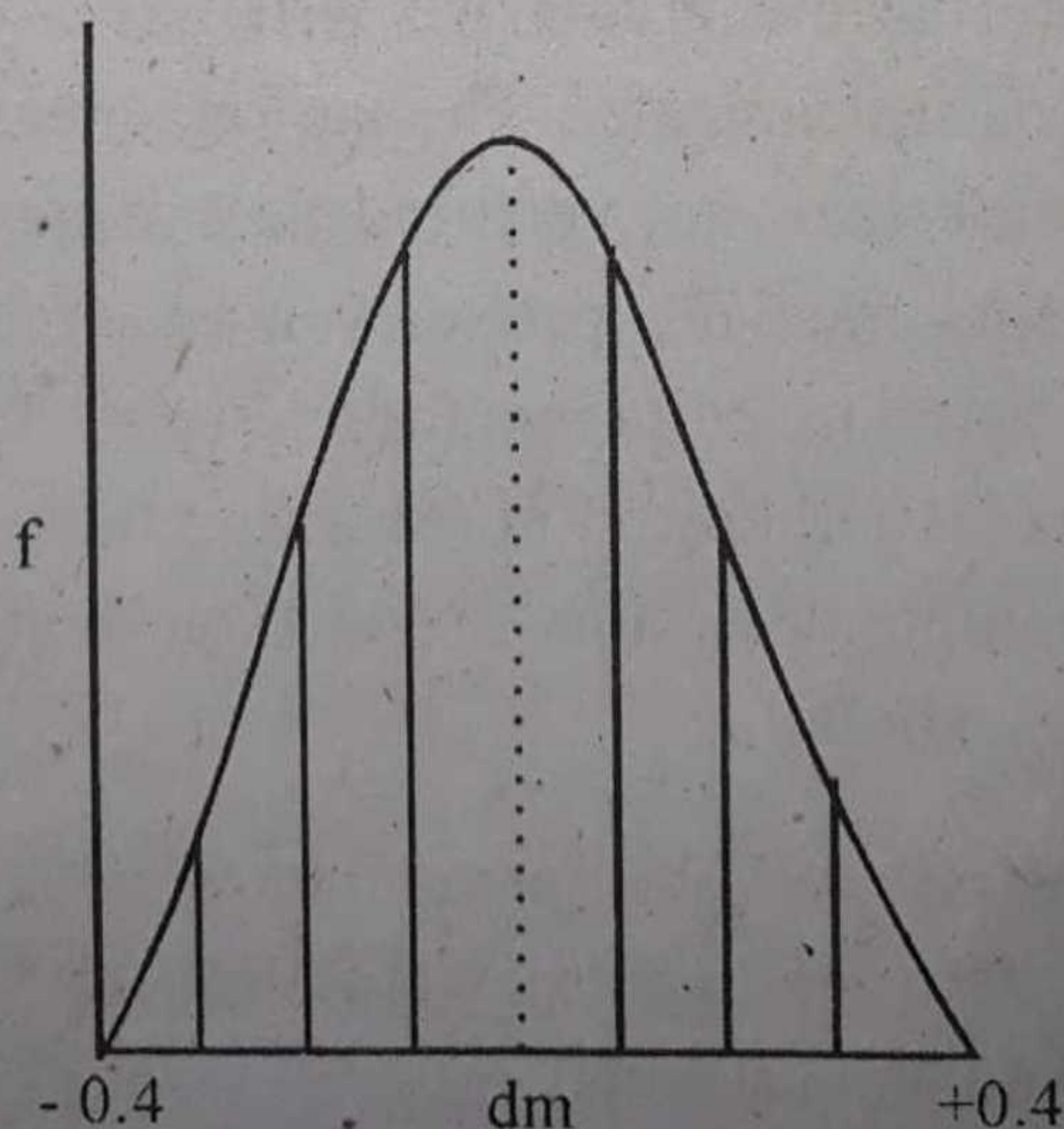
In chemical analysis, indeterminate errors follow this Gaussian type distribution. For e.g., if the deviation from the mean (dm) of hundreds of repetitive weighing measurements on a single object, are f plotted against the frequency (f) of occurrence of each deviation we get a curve a shown in figure. This curve proves the fact that a number of small independent and uncontrolled uncertainities are there in our normal measurements.

- 0.4   0   +0.4

dm
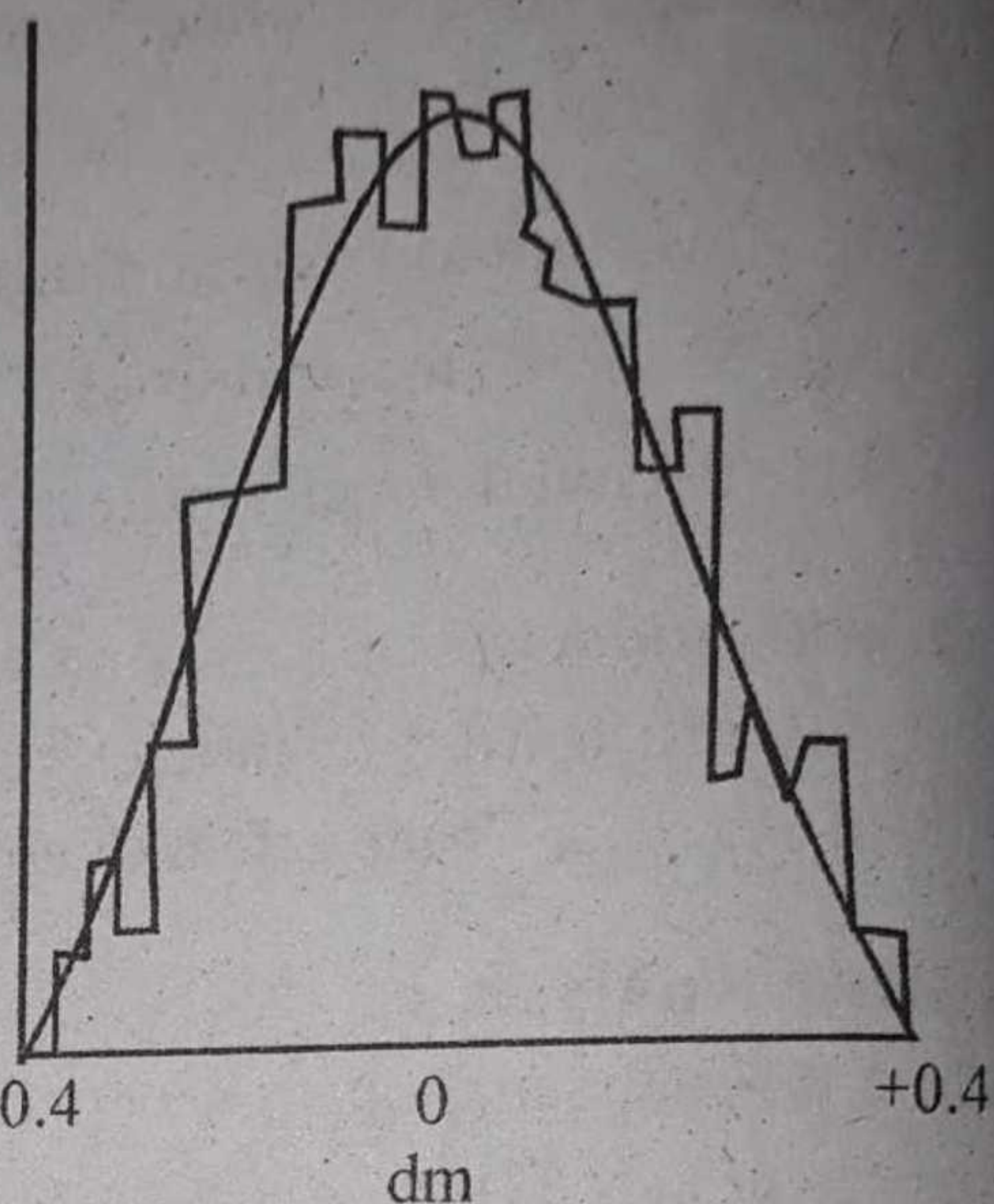
These uncertainities manifest in the result.

This Gaussian distribution of most analytical data, permits us to use statistical techniques to estimate the extent of indeterminate errors. Thus we use several statistical techniques like mean, median, average deviation, standard deviation etc., while analysing our experimental results.

Minimising Errors :

From the above discussion it is clear that, if we want to minimise error we have to minimise determinate as well as indeterminate errors.

To minimise determinate errors we will have to use standard, internationally accepted instruments. The various measuring aids must be calibrated periodically and got certified according to international standards. All reagents must be properly maintained. Dependable procedures must be adopted to avoid methodic errors. To avoid personal errors one must be careful and honest in recording the observation. It is human tendency to manipulate results to get high degree of precision. This tendency must be avoided.

We know that indeterminate errors are uncontrollable. So, to minimise these errors, we repeat the experiments several times and adopt statistical techniques to get maximum precision.

The following are some other specific suggestions to minimise error.

1. Blanks experiments are to be conducted along with regular ones.

2. Atmost care is taken to avoid personal errors. Arithmatic mistakes, etc., should not be committed.

3. While taking readings one must be very careful to note the correct reading.

4. Without getting proper and complete instructions, the experiment should not be done.

5. When one becomes tired, the experiment must be stopped in a convenient place and continued after taking sufficient rest.

2. PRECISION

Definition :

It is the degree of aggrement between two or more measured values of a property measured under identical condition.

Explanation with example :

Let the weight of a beaker be 20.0326g when weighed ina particular set of conditions. If the same beaker is weighed under identical conditions and if the weight got is the same 20.0326g then we say that there is precision in the weighing. Thus if a value is reproducible then it called a precise value.

Precision and range of a set of value :

The differenses between the highest and lowest values in a set of values is called the range (w). This range is a measure of precision. If the range is wids then it means that the measurement is less precise.

3. ACCURACY

Definition :

It is the degree agreement between the measured value and the expected or true value of a property.

Explanation with example :

As per analytical data available the solubility product of AgCl is 1.8×10^{-10}. If an analyst gets this same value while determining the same

in an experiment, then. It is said accuracy of the result is excellent. If it differs, we calculate the deviation. This deviation gives a measure of the accuracy of experiment.

Methods of expressing accuracy :

Accuracy is expressed in terms of absoulte error or relative error. The lower these values are, the more will be the accuracy.

Absolute Error (E) :

Definition :

It is defined as the difference between the accepted value (x_t) and the observed values (x_i). Mathematical expression : $E = x_i - x_t$.

Relative Error (RE) :

Defnition :

It is the error percentage the accepted value.

Mathematical expression :

$$RE = \frac{E}{x_t} \times 100$$

Advantage of relative error over absolute error :

Absolute error depends onthe reliability of the accepted value it self as a lot of uncertainity may be there about the accepted value. So relative error is used more often to express accuracy.

Distinctions between Absolute Error (E) and Relative Error (RE)

		E	R E
1.	Definition	$E = x_i - x_t$	$RE = (E/x_i) \times 100$ $= (x_i - x_t / x_i) \times 100$
2.	Usefulness	Not much useful as E depends on x_t which itself is subject to uncertainty.	More useful

Distinctions between Precision and Accuracy

	Precision	Accuracy
1.	Degree of agreement between a value and other values obtained under substantially the same condition.	Degree of agreement between a measured value and the true value or expected value
2.	Precision may be achieved.	Accuracy is never known. It is known within certain limits only. Accuracy can be approached but never attained.
3.	Expresses the reproducibility of the results.	Expresses correctness of a measurement

One must clearly understand the connotations of the terms precision and accuracy. An understanding of the two terms will make it clear that high precision does not imply accuracy. We can prove this with an example. In a volumetric estimation one may get concordant titre values. Yet the result may turn out to be wrong. Here, as concordant titre values have been got, the result is precise.

But as the answer doesnot agree with the expected value, it is not accurate. Thus as precise value may not be an accurate value. The reason for this may be a determinate error like instrumental error or operative or personal error like personal carelessness or even some unknown indeterminate error.

Confidence limits :

Definition :

They are the limits, which may be set, about the experimentally measured mean (x), with in which we may expect to find the true mean (μ) with a given degre of proability.

Explanation :

Only the mean of an infinite number of measurements can be true or accurate mean. This true mean is represented by μ. This is unattainable as an infinite number of measurements is impossible. i practice, we make a finite number of measurements and calculate the experimental mean ($\bar{x}$) How near or how far away is $\bar{x}$ from μ ?. This can be determined by setting limits. Within these limits we can find μ with a given degree of probability. The interval between these limits is called the confidence interval. The size of this interval depends on the degree of probability that we need. If we want 99% probability, the intervel will be small 99% probability means that 99 times out of 100, the true mean will be within this intervel. Similarly we can choose 95% probability, 90% probability etc., depending upon our needs. This percentage probability is called confidence level.

Mathematical expression of confidence limits :

a. When the standard deviation is for a small number of measurements, σ (standard deviation, applicable only when the number of measurements is large) the confidence limit for a single measurement for μ is given by

$$\mu = \bar{x} \pm Z\sigma$$

Where $Z = \dfrac{\bar{x} - \mu}{\sigma}$

The confidence limit for the μ of N measurements is given by

$$\mu = \bar{x} \pm \dfrac{Z\sigma}{\sqrt{N}}$$

When σ is unknown : Confidence limit for μ of N measurements is given by

$$\mu = \bar{x} \pm \dfrac{ts}{\sqrt{N}}$$

Where
$$t = \frac{\overline{x} - \mu}{s}$$

Thus we find by applying statistical methods we can fix the confidence limits within which the true average of a set of experimental results can be found for various confidence levels of probabilities. For this we must known the values of Z and t which are readily available in literature.

Rejection of Results :

In a set of data we come across one or two values that are suspicious. Whether to reject such a data or not would become difficult. In such cases we employ a test called Q - test. In this test we compare two Q - values.

i. Q_{exp} and ii. Q_{crit} If $Q_{exp} > Q_{crit}$ then we reject the data. If not, we retain the data.

To get Q_{exp} the difference between the suspicious value and its nearest neighbour is divided by the spread of the entire set.

To get Q_{crit} the following table giving various Q_{crit} values is used.

Number of observations	Q_{crit}		
	90% confidence	96% confidence	99% confidence
3	0.94	0.98	0.99
4	0.76	0.85	0.93
5	0.64	0.73	0.82
6	0.56	0.64	0.74
7	0.51	0.59	0.68
8	0.47	0.54	0.63
9	0.44	0.51	0.60
10	0.41	0.48	0.57

Let us illustrate how to decide whether to reject or retain a data.

Let us say that the experimentally determined percentages of CaO in a calcite sample are 55.95, 56.00, 56.04, 56.08, 56.23. Now we suspect that the last value and so we have to decide whether to reject it or not.

For this set of data Q_{exp} = (Suspecious value – Its nearest neighbour) ÷ the spread of the entire set.

$$= (56.23 - 56.08) \div (56.23 - 55.95)$$

$$= 0.15 \div 0.28 = 0.54$$

Now let us say we want our results in 90% confidence level i.e., our result should be 90% as near to the correct value then we look under 90% confidence column of the table. Since we have 5 observations we look for Q_{crit} against number of observations 5 and under 90% confidence. The value is 0.64.

Since $Q_{exp} < Q_{crit}$ we conclude that the value is to be retained.

Though the Q - test is superior to other methods one must be careful while deciding to either reject or retain a particular data using this test. One has to use one's good judgement in deciding. Where the number of observations are small the reliability of Q - test decreases. So, to reject a value in a small set a cautious approach is desirable.

4. SIGNIFICANT FIGURES

They are figures in a number which contains only digits known with certainty plus the first uncertain one.

Explanation :

A measured value has some uncertainty about it. There is a convention to give the measured value as a number such that it contains only one figure about which there is uncertainty. The practice is called significant figure convention.

Example :

If the weight in a weighing is known with certainty only upto three decimals, the value should be reported only upto four decimals.

Salient features ? Points to be borns in mind while using the concept of significant figures :

The number of significant figures in a given number is found out by counting the number of figurs from left to right in the number begining with the first non - zero digits and continuing till reaching the digit that contains the uncertainty.

Examples :

1. Each of the following has three significant figures. 583. 0.234, 1.67, 0.00987 and 65.4.

2. Zero is a siginificant figure when used as a number. It is not a significant figure when it is used to locate decimal points in very small and very large numbers. E.g., 0.02670 has four siginificant figures. The two zeros before 2 are used to imply only the magnitude. So they are not significant. The zero beyond 7 is significant.

3. The value 6.030×10^{-1} has four significant figures while 1.45×10^5 has three significant figures.

Using the above, one can determine the number of significant figures in a given number.

Its importance or Uses :

In presenting scientific data, one comes across a set of values. For this set of values, one gives their mean or median as the best value. Now the uncertainty about this best value must also be indicated while presenting the data. To achieve this, usage of significant figures is very helpful. On several occasions one has to round off numbers to give meaningful results. For this, the practise of rounding off to a numer which contains only one uncertain figure in it, is adopted. Thus significant figures become important.

Methods of expressing Precision :
Precision is expressed by two methods.
i. Absolute Method
ii. Relative Method.

i. Absoulte Method :

In this method the precision is expressed in terms of average mean deviation. The smaller the value of average mean deviation the greater will be the precision.

ii. Relative Method :

In this method the precision is expressed in terms of percentage deviation from the mean of a set of values.

$$\left.\begin{array}{l} \text{i.e., Percentage deviation} \\ \text{from the mean} \end{array}\right\} = \dfrac{\text{Mean of a set of values}}{\text{minus a particular value}} \times 100$$

$$= \dfrac{\overline{x} - x_i}{\overline{x}} \times 100$$

To understand the above two methods let us learn some terms used in statistics.

Mean / Artithemetic mean / Average :

Mean is the quotient obtained when the sum of a set of replicate measurements by the number of individual results in the set. The following worked out unversity problem would illustrate mean.

Problem :

Calculate the mean for the following set of values 20.21, 20.04, 20.13 and 20.19

Solution :

Mean = $(20.21 + 20.04 + 20.13 + 20.19) \div 4 = 80.57 \div 4 = 20.14$.

Median :

It is the value about which all other values are equally distributed. Half the values will be greater and the other half smaller than the median.

The median is obtained by arranging the set of values in increasing or decreasing order. If there are odd number of values then the middle

value gives the median. If there are even number of values than the average of the middle pair gives the median. The following worked out university problems would illustrate the median.

Problem :

Calculate the median for 20.20, 20.08 and 20.02.

Solution :

Arranging the values in increasing order we get 20.01, 20.08 and 20.20. This set contains odd number of values. Therefore the median of this set is the middle value i.e., 20.08.

Problem :

Calculate the median for the following set of values 20.21, 20.04, 20.13 and 20.19.

Solution :

Arranging the values in increasing order we get 20.04, 20.13, 20.19 and 20.21. This set contains even number of values. Therefore the median of this set is the average of the middle pair of values, i.e., (20.13 + 20.19) ÷ 2 = 40.32 ÷ 2 = 20.16

Distinctions between Mean and Median

		Mean	Median
1.	Definition	It is the quotient obtained when the sum of a set of replicate measurements by the number of individual results in the set.	It is the value about which all other values are equally distributed 20.08
2.	Examples	20.09 Date : 20.20 20.08, 20.01	20.08
3.	Precission	Better than median	Less than mean

Mean deviation or Average deviation

Definition :

The average deviation of a value in a set of values is the average of the divations of all the individual values from their average.

Explanation :

To get the average deviation.

i. The average of the given set of values is calculated.

ii. The deviation of each value from the average is calculated.

iii. The average of all these deviations (ignoring signs) give the average deviation. The following worked out University problems would illustrate average deviation.

Problem :

Calculate the mean (average) deviation for the following set of values : 20.21, 20.04, 20.13 and 20.19.

Solution :

Average of this set $= (20.21 + 20.04 + 20.13 + 20.19) \div 4$

$$= 80.57 \div 4 = 20.14.$$

Value	Deviation from average
20.21	$20.14 - 20.21 = 0.07$
20.04	$20.14 - 20.04 = 0.10$
20.13	$20.14 - 20.13 = 0.01$
20.19	$20.14 - 20.19 = 0.05$

Total of deviation from average $= 0.07 + 0.10 + 0.01 + 0.05$

$$= 0.23$$

Mean deviation $= 0.23 \div 4 = 0.575$

Its usefulness :

If the mean deviation of a set of measure ments is small it means that the average of that set is nearly precise.

Standard Deviation :

Definition :

It is the square root of the quotient obtained by dividing the sum of the squares of the individual deviations from their mean by the number of measurements made.

or

Standard deviations $=\sqrt{\begin{array}{l}\text{Sum of squares of individual deviation from}\\\text{their mean} \div \text{Number of measurements made}\end{array}}$

Explanation :

To obtain standard deviation.

i.   The average $(\bar{x})$ of the measurements (x_{ii}) is calculated.

ii. The individual deviation of each measurement from the average $(x_i - \bar{x})$ is calculated.

iii. Each individual deviations is squared $(x_i - \bar{x})^2$

iv. All the individual deviation square are added $\sqrt{\Sigma(x_i - \bar{x})^2}$.

v.   The value obtained in step.

vi. It is divided by the number of measurements made $\Sigma(x_i - \bar{x})^2 \div N$.

vii. The square root of the value obtained in step.

v.   Gives the standard deviation.   Thus standard deviation $\sigma = \sqrt{\Sigma(x_i - \bar{x})^2 \div N}$.

σ is applicable only when the number of measurements is large. But in analytical chemistry we make only a small number of measurements. so in step (v) instead of dividing by N the value obtained in step (iv) is divided by N $-$ 1.   The standard deviation for a small number of measurements $s = \sqrt{\Sigma(x_i - \bar{x})^2 \div (N - 1)}$.

Example / Illustration :

The following worked out university problem would illustrate how standard deviations is obtained from a set of data.

Problem :

Find the standard deviation for a subset having the following six values.

7.720, 7.725, 7.736, 7.719, 7.742 and 7.751

Solution :

i. Calculation of average $\bar{x}$

$$(7.720 + 7.725 + 7.736 + 7.719 + 7.742 + 7.751) \div 6$$

$$= 46.393 \div 6 = 7.732$$

ii. Calcualtion of $\Sigma(x_i - \bar{x})^2$

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
7.720	0.012	1.44×10^{-4}
7.725	0.007	0.49×10^{-4}
7.736	0.004	0.16×10^{-4}
7.719	0.013	1.69×10^{-4}
7.742	0.010	1.00×10^{-4}
7.775	0.043	18.49×10^{-4}

iii. $\therefore \Sigma(x_i - \bar{x})^2 = 23.37 \times 10^{-4}$

iv. $\dfrac{\Sigma(x_i - \bar{x})^2}{N - 1} = \dfrac{23.27 \times 10^{-4}}{5} = \dfrac{4.82}{5} = 0.96$

v. $\sqrt{\Sigma(x_i - \bar{x})^2 \div (N - 1)} = 0.98$

Standard deviation $= 0.98$.

Exercises : (Univeristy Problems)

1. Calculate the standard deviation for the subset of the values : 25.32, 25.07, 25.18, 25.26
   [Ans : 0.1082]

2. The percentage of a constituent M in a compound MA are found to be 62.42, 62.28, 62.46, 62.32 and 62.22. Calculate the mean, median and standard deviation.
   [Ans : i. 62.34, ii. 62.32, iii. 0.099]

Advantages of standard deviation : Standard deviation is more reliable than average or mean deviation to express precision as it has theoretical foundation.

Its usefulness :

If the standard deviation of a set of measurements is small, it means that the average of the set is nearly precise.

### Distinctions between Mean deviation and Standard deviation

		Mean deviation	Standard deviation
1.	Definition	It is the average of the deivation of all the individual values from their average.	Sum of squares of individual deviation from their mean ÷ number of meansurements made.
2.	Formula	$\sqrt{\Sigma(x_i - \overline{x})} \div N$	$\sqrt{\Sigma(x_i - \overline{x})^2} \div (N-1)$
3.	Precision	Less than standard deviation but better than average	Better than mean deviation, It is the best way of expressing precission.

Curve Fitting - Method of Least Squares

If we want to present a trend or relationship we draw a graph. A graph is obtained by plotting two variables X and Y against each other. Let us consider the following data.

x	y
2	20.20
4	40.4
6	60.6
8	80.2
10	100.8
12	120.0

A graph may be drawn using these data (figure). By convention we plot the independent variable, i.e., the cause of the value based on which a particular prediction is made along the axis (horizontal axis) and the dependent variable. i.e., the effect on the predicted property, along the Y axis (vertical axis).
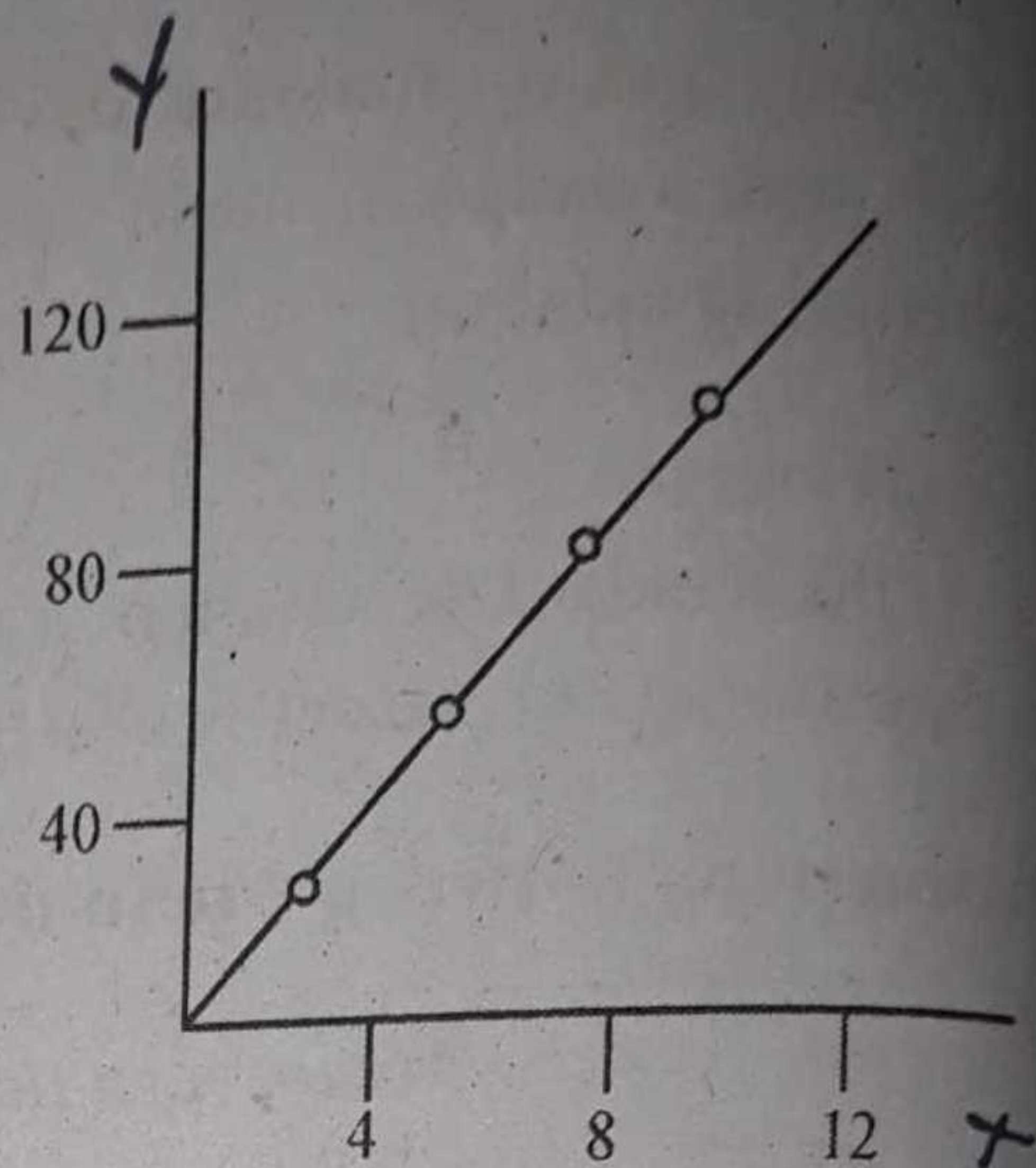
The graph gives the trend or relationship between X and Y. The above graph indicates a linear relationship between X and Y. In the above graph linearity is good. Usually for the data obtained in experiments, linearity will not be as good as shown in figure. In such cases we have to draw a best fit line. This is called curve fitting. For this purpose we use the method of least squares. By this method we get a straight line for which the algebraic sum of the vertical deviations is less then that from any other straight line. There will be only one such line i.e., it will be the best fit line. This best fit line cannot be drawn just by observing the points on the graph. For drawing the best fit line we have to use the method of least squares.

Let us consider the following data :

X	Y
1	1.5
2	1.8
3	2.7
4	4.0

To draw the best fit line for the above data the following steps are followed.

i. ΣX, ΣY, ΣXY and ΣX^2 are calculated.

In our example,

$\Sigma X = 10$, $\Sigma Y = 10$, $\Sigma XY = 29.2$ and $\Sigma X^2 = 30$

ii. The above values are substituted in the following simultaneous equations.

$$\Sigma Y = aN + b\Sigma X$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

Here N = the number of pairs of data = 4

i.e., $10 = a \times 4 + b \times 10$ - (1)

and $29.2 = a \times 10 + b \times 30$ - (2)

multiplying (1) by 3 we have

$30 = 12a + 30b$ - (3)

substracting (2) from the (3) we have

$0.8 = 2a$

$\therefore a = 0.4$ - (4)

Substituting (4) in (1) we have $b = 0.84$

iii. Now for the given values of X, Y values are calculated using the equation

$$Y_{cal} = a + bX$$

The calculated Y values for two values of X from our data are calculated.

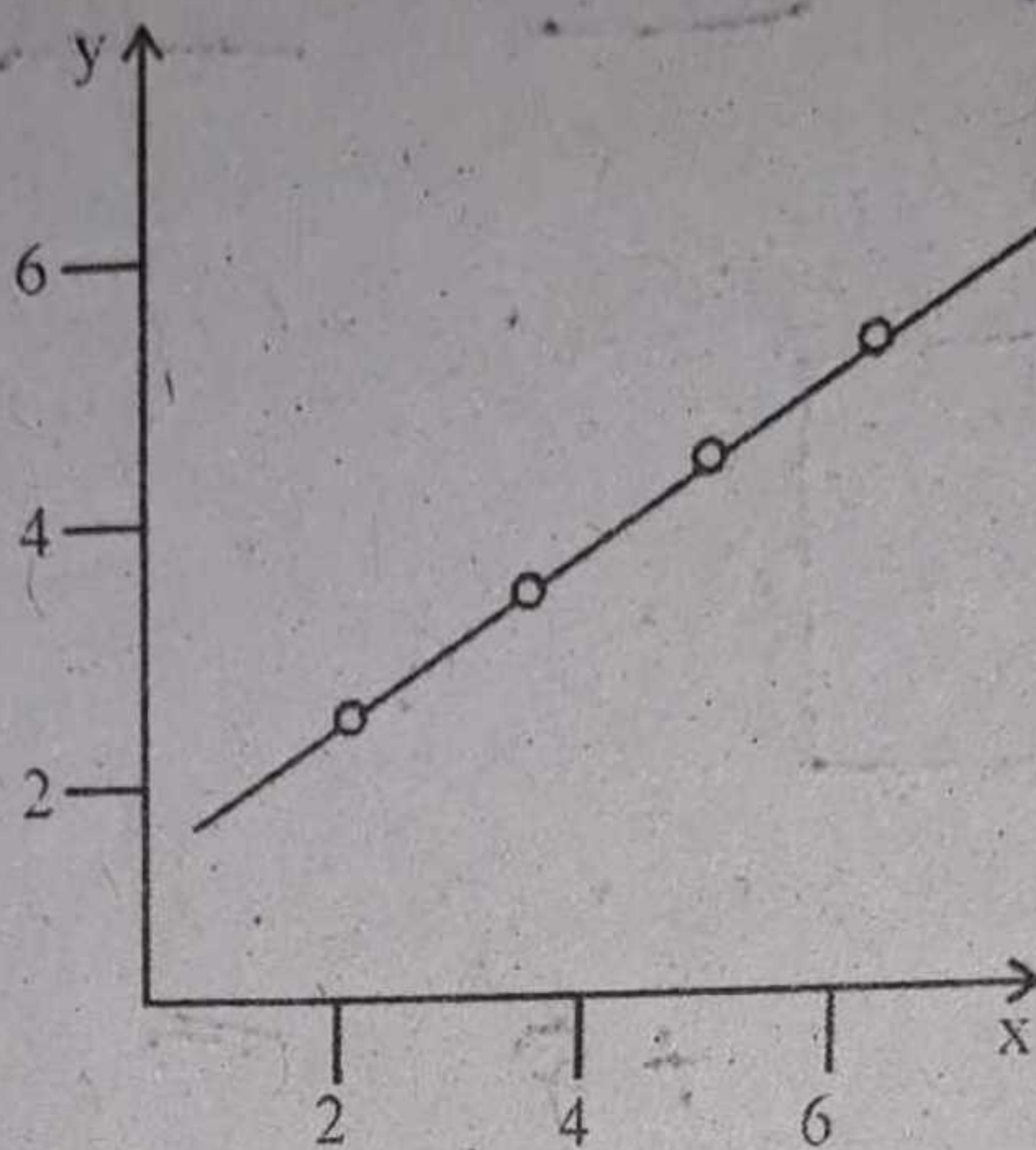| X	$Y_{cal} = (0.4 + 0.8) X$
1	1.2
4	3.6

iv. Two points corresponding to these two pairs of values are marked on the graph and they are joined. We get the best fit line for our data. The following will be the graph for the data given.

Correlation Coefficient :

It is a quantity that indicated the extend of linearity of a given set of data. It is denoted by r.

Mathematical expression.

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\{[N\Sigma X^2 - (\Sigma X)^2][N\Sigma Y^2 - (\Sigma Y)^2]\}^{\frac{1}{2}}}$$

Where X and Y are vaiables or two properties, N is the number of pairs of data.

Explanation :

It helps us find out whether there is any linear relationship between two properties X and Y from N pairs of data relating X and Y.

If r is 1.0 is means that there is perfect linearity relationship between X and Y.

If r is between 0.99 and 0.75 it means that there is excellent linearity relationship between X and Y.

If r is between 0.75 to 0.50. If means that their is good linearity realtionship between X and Y.

If it is less than 0.50 we can conclude that there is poor linearity relationship between X and Y.

In our example given above

$N = 4$ $\Sigma X = 10$ $\Sigma Y = 10$

$\Sigma XY = 29.2$ $\Sigma X^2 = 30$ $(\Sigma X)^2 = 100$

$\Sigma Y^2 = 28.78$ $(\Sigma Y)^2 = 100$

$$r = \frac{(4 \times 29.2) - (10 \times 10)}{\{[4 \times 30 - 100][4 \times 28.78 - 100]\}^{\frac{1}{2}}}$$

$$= \frac{116.8 - 100}{\{[120 - 100][115.12 - 100]^{\frac{1}{2}}}$$

$$= \frac{16.8}{\{[20][15.12]\}^{\frac{1}{2}}} = \frac{16.8}{304^{\frac{1}{2}}}$$

$$= \frac{16.8}{17.44} = 0.96$$

That is, as per our data, there is excellent linear relationship between X and Y.

Thus correlation co - efficient givens us and idea about the extent of linearity of a given set of data.

Unversity Questions

1. Give two examples corrosive chemicals.
2. Explain the methods used of the storage and handling of corrosive reagents.
3. Mention an example for a flammable chemical.
4. How are flammable chemicals stored?
5. How will you handle a flamable chemicals.