UNIT 16

EMPIRICAL DATA COLLECTION AND ANALYSIS

Student Learning Outcomes (SLOs)

- Differentiate between qualitative data and quantitative data (qualitative data includes all non-numerical information obtained from observation, not from measurement. Quantitative data are obtained from measurements and always associated with random error/uncertainties determined by use of apparatus and the human limitations such as reaction times)
- Justify the propagation of random errors in data processing shows the impact of the uncertainties on the final result. Some examples main include;
  *When we process data that contains random errors these as errors can propagate or accumulate throughout the calculation resulting in large uncertainties in the final result.
- For example if we measure the length and width of a rectangle to calculate its area, any small error in the measurement of length and width will propagate through the area calculation resulting in larger uncertainties in the final area measurement. This information is critical in the scientific research as it helps us assess the reliability of our data and draw valid conclusion from our experiments.
- Analyse the concept that experimental design and procedure usually lead to systematic errors in measurements, which causes are deviation in a particular direction.
- Justify that repeat trials and measurements will reduce random errors but not systematic errors.
- Explain that graphical techniques are an effective means of communicating the effect of an independent variable on a dependent variable and can lead to the determination of physical quantities.
- Discuss that sketched graphs have labelled but unscaled axes, and are used to show qualitative trends such as variables that are proportional or inversely proportional.
- Discuss that drawn graphs have labelled and scaled axes, and are used in quantitative measurements.

Analytical chemistry is mainly related to the chemical characterization of materials. Everything we use or consume is made up of chemicals and knowledge of the chemical composition of any substance is important in our daily life. Analytical chemistry plays an important role in nearly all aspects of chemistry, for example, agricultural, clinical, environmental, forensic, manufacturing, metallurgical, and pharmaceutical chemistry.

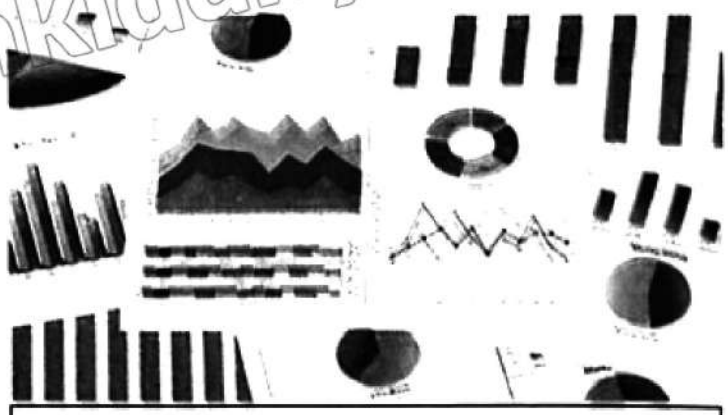


Fig 16.1

The percentage of nitrogen in fertilizers determines its value. The food quality is determined by the presence of essential nutrients, minerals and vitamins in it. The air quality is determined by monitoring the pollutants in it, which contaminate the air. In diabetics, blood glucose is monitored periodically. The presence of trace elements from gunpowder on a murder defendant's hand will prove a gun was fired. The efficacy of drugs is determined by their purity. Thus in every field of life analytical chemistry is involved in one way or the other.

16.1 Types of Data

Chemical characterisation data is of two types. Both are important for understanding chemical phenomena.

1. Qualitative data
2. Quantitative data

Qualitative data in chemistry refers to non-numeric information derived from observations about chemical characteristics and reactions. For example; observing the colour change of a reagent in a solution to determine the presence of specific ions or molecules, identifying the type of chemical reaction (e.g., exothermic, endothermic, or absorption), reporting the chemical sample's odour, colour and state (solid, liquid or gas). This means qualitative data is about describing the properties and behaviour of substances and reactions.

On the other hand, quantitative data refers to numerical measurements that are obtained from experiments. Quantitative data is obtained from instruments. For example concentration, weight, volume, temperature, etc. This means quantitative data, is about measuring and calculating certain numerical values.

Quantitative data are often obtained from measurements of substances and reactions. These measurements can include mass, volume, concentration, temperature, and other properties. Random error and uncertainty are always present in these measurements, which arise from the limitations of the apparatus used and human factors. For examples;

(1) For determining concentration of a substance burette, pipette, and volumetric flask is used for titrations. Inaccurate calibration of these glass wares, improper cleaning of glass ware, and human error in reading the meniscus are sources of errors. In a

titration experiment to determine the concentration of an acid, the volume of the titrant added from a burette is measured. Small errors in reading the meniscus level or the initial and final readings can lead to uncertainties in the calculated concentration.

(2) For mass measurement an analytical balance is used. Improper calibration of the balance, air currents in the laboratory, and cleanliness of the balance pan can cause error in measurement. Weighing a sample of a chemical compound for a reaction, even a slight draft in the lab or dust on the balance pan can introduce errors in the measurement.

16.2 Types of Errors

Error is the difference between the value or quantity obtained in an experiment and the value accepted in the experiment or in the literature. Primarily there are two types of errors in experiments.

(1) Random errors

Random error is the random difference between the observed value and the true value.

Random errors can be affected by:

(a) How easy the instrument or scale is to touch and read. A person reads the scale incorrectly.
(b) Changes in the environment, such as temperature fluctuations in the laboratory air currents in the room, etc.

Random errors will cause the result to deviate from the accepted value in either direction (either too high or too low). Repeating the experiment and working with the average of the results can help reduce the impact of random errors

(2) Systematic errors

Systematic error is a consistent repeatable error due to faulty equipment flawed experimental or human mistakes.

Systematic errors always pull the result away from the accepted value in the same direction (always too high or always too low).

For example, For instance, if an electronic balance is not zeroed (using a tare button), the mass weighing are always greater than they should be. If a volume is not read at eye level in a burette, the volumes are always lower than they should be because of a parallax (parabolic) error. If a cap is not kept on a spirit burner during calorimetry experiments, the alcohol evaporates, resulting in a larger mass loss. Repeating the test and working with the mean will not eliminate systematic errors.

Measurement techniques are used.

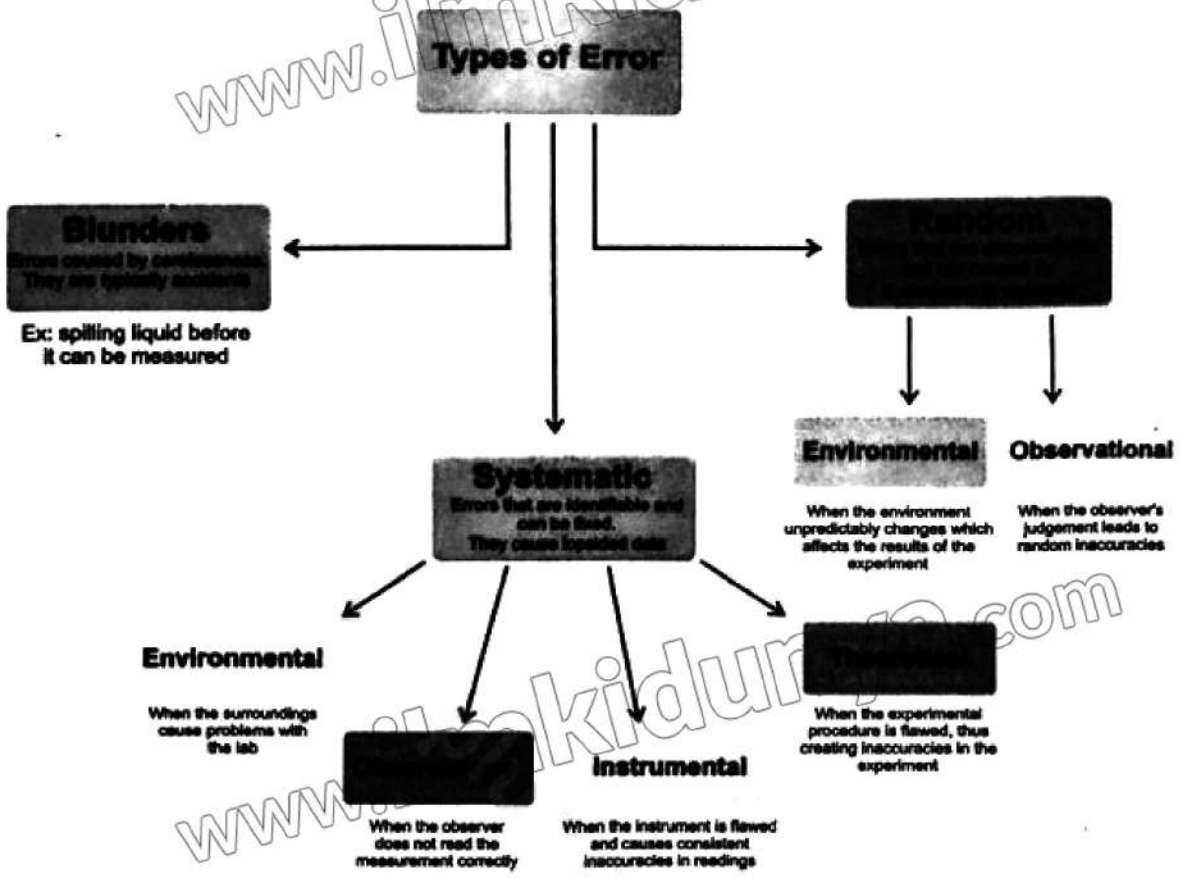


Fig 16.2

16.2.1 Propagation of random errors

When we process data that contains random errors, these errors can propagrate or accumulate throughout the calculations, resulting in larger uncertainties on the final result.

For example if we measure the length and width of a rectangle to calculate its area, any small error in the measurement of length and width will propagate through the area calculation, resulting in larger uncertainties in the final area measurement.

Let us consider an example: In the scale, the scale has division of 1cm, so the measurement can be nearest to 0.1cm

Length of the rectangle (L) = 10.0 cm

Uncertainty in length measurement (ΔL) = 0.1 cm

Width of the rectangle (W) = 5.0 cm

Uncertainty in width measurement (ΔW) =0.1 cm

Area of rectangle = Length x Width

= 10 cm x 5 cm

= 50 cm^2

Propagation of uncertainty:

The uncertainty in the area (Δ) an be calculated as follows;

$$\Delta A = [W \times \Delta W] + [L \times \Delta L]$$

$$= [5.0 \text{ cm} \times 0.1 \text{ cm}] + [10.0 \text{ cm} \times 0.1 \text{ cm}]$$

$$= [0.5 \text{ cm}^2] + [1.0 \text{cm}^2]$$

$$= 1.5 \text{ cm}^2$$

Calculated area of the rectangle is 50.0 cm^2

The uncertainty in the area is 1.5 cm^2

So, the final area of the rectangle can be reported as 50 ± 1.5 cm^2. This means due to uncertainty in the measured length and width, the area of the rectangle could be as low as 48.5 cm^2 or as high as 51.5 cm^2. This range of possible values allows us to understand the practical variability in our measurements. This information is critical in the scientific research as it helps us assess the reliability of our data and draw valid conclusion from our experiments.

16.2.2 Systematic errors and 'experimental design and procedure

The design and implementation of experiments are carefully planned to reduce errors, but they can inadvertently introduce systematic errors that consistently skew measurements away from the true value. These errors can be caused by various factors, including equipment calibration, environmental conditions, and human biases, leading to a decrease in the accuracy and reliability of experimental results.

Example1:

You have designed an experiment to determine the concentration of a Alizarin or any other compound in a solution using a spectrophotometer. The following systematic errors can occur:

Error in calibration: Before using the spectrophotometer, it needs to be calibrated with standard solutions of known concentrations. After that calibration curve is developed. If there's an error in the calibration process, such as incorrect standard solution preparation or inaccuracies in the calibration curve. All final results would be doubtful.

Sample Preparation: The way you prepare the sample can bring in systematic errors. For example, if you consistently add slightly more or less of the sample solution to the cuvette each time, it will increase the concentration and in turn the measured absorbance and thus the concentration.

Instrument performance: Over time, the spectrophotometer might drift in its performance, leading to consistent errors in absorbance (optical density) measurements. This could be due to factors like aging of the instrument's components or damage due to shifting and misuse.

Impurity effect:

The impurity effect can also be called matrix effect. In the matrix if the impurities are present other than your compound of interest, they can interfere in the optical density or light absorption intensity, giving rise an error in the concentration measurements.

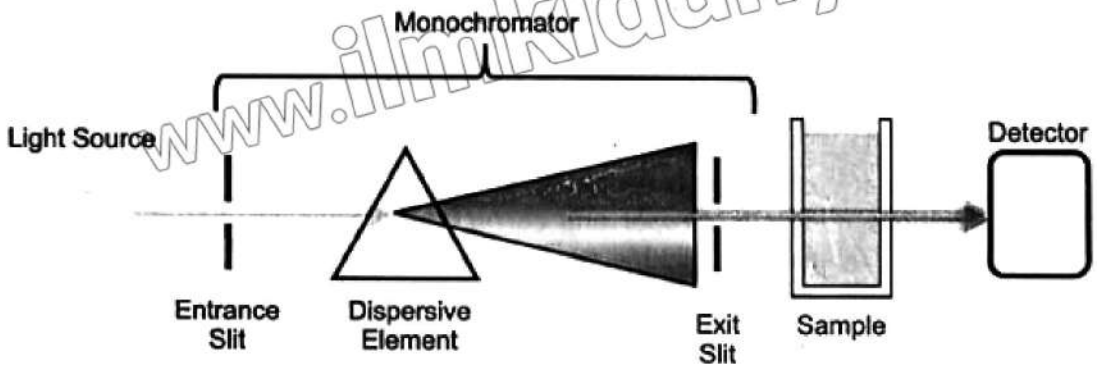


Fig 16.3

Example 2: If an analyst or student tries to find out the concentration of a reducing agent ($FeSO_4$) solution with the help of acidified $KMnO_4$ volumetrically, the following systematic errors can occur:

i) Error in volume measurements:

If a student measures the volume without properly calibrated burette pippet, or if the meniscus reading is inaccurate this will cause a deviation from the true value in the results.

ii) Impurity in analytical reagents:

If the reagents used like $KMnO_4$ or sulphuric acid has a degree of impurity in them the measured concentration will have a deviation from the true value of the sample concentration.

iii) Effect of temperature fluctuation:

The rate of the reaction and equilibrium is greatly affected by temperature changes which may lead to a potential error in the titration. For example, the rate of the reaction may increase by warming mixture of sulphuric acid and ferrous sulphate during titration, which leads to the deviation of concentration measurements from true value.

iv) Incomplete redox reaction:

If the redox reaction between oxidizing and reducing agents is not complete, it can lead to incorrect endpoint and consequently, the concentration calculations will become incorrect. This may be due to inhibitors present in the reaction mixture or due to inadequate stirring.

To avoid such systematic and random errors in the experiment, we should revise the experimental design and procedure.

16.3 Repeat trials and the elimination of errors

The random errors can be reduced by repeat trials and measurements in chemical analysis, as they help assess uncertainty or variability in the analysis. This assures a more accurate determination of the desired quantity. This is due to the fact that random errors are typically unpredictable, and variations usually occur in different readings. It makes it necessary to take repeated and concordant readings to estimate associated uncertainty, along with true value.

However, systematic errors cannot be eliminated by repeat trials, because systematic errors are consistent. These are determinate errors means they are predictable errors, and they

happen due to certain known reasons. Systematic errors are caused by reasons such as incorrect calibration of instruments, impurity in analytical reagents, or due to temperature fluctuations. These reasons deviate the measurements in the same fashion either positive or negative.

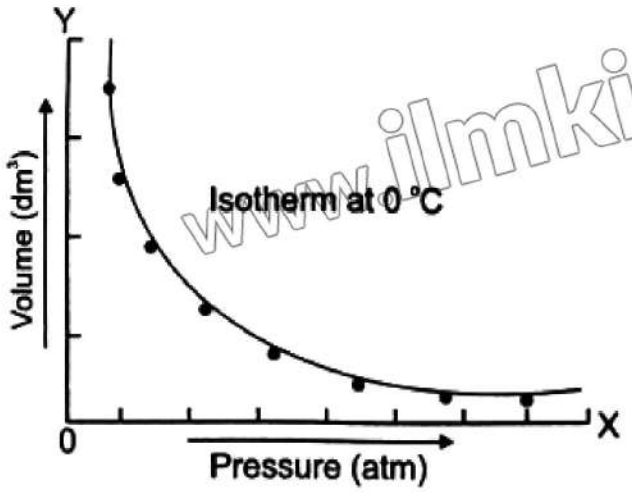
To rule out the systematic errors, it is necessary to control them after identifying the source of error. This can be achieved by proper and correct calibration of instruments, using high quality and the purest reagents available, and ensuring consistent conditions in the lab. This practice will minimise the impact of systematic errors on accuracy.

16.4 Graphical techniques and data of variables

Drawing graphs is an effective method to represent the relationship between two variables i.e. the effect of independent variable on a dependent variable.

Dependant variable: The variable that is affected by changes in the independent variable and this is usually plotted on y - axis.

Independent variable: It is manipulated or controlled by the individual who is doing experiment. It is plotted on x - axis.



Graph 16.4

Fig 16.5

For example, in the pressure-volume graph (Boyle's law), pressure is independent variable and the volume is the dependent variable.

By making a graph between two variables the relationship is visualised, and visually comprehended that how does an independent variable affects on the dependent variable. The graphs depict the relationship between variables, whether they are direct, inverse or more complex.

The graph can also be used to determine certain physical quantities, for example, the rate of change (slope = $\tan\theta$) between the variables. The slope can represent important physical properties like acceleration, rate of reaction, or consistency in a relationship.

Thus, the graphical method of representing data has greater visual appeal than a table and also represents relationships which may help in drawing extra plots.

16.5 Sketched Graphs

16.5.1 Sketched Graphs; Unscaled but Labelled Axes

Sketched graph is a way of visual representation, whose purpose is to represent qualitative trends rather than precise mathematical values. These graphs have the following salient features.

Labelled Axes

The variables which are to be represented on x and y axis are labelled in the sketch graphs, but numerical values are missing in these graphs, i.e. they are not scaled. For example the in the graph 16.1 which is a sketched graph, the axes are labelled as volume and pressure but there is no value or units of pressure and volume along the axes.

Unscaled Axes

Since the graphs are unscaled there is no grid line or tick marks which correspond to specific numerical increments. The major focus of this graph is shape and direction of the line or the curve.

Uses and Advantages

Sketched graphs serve various purposes effectively:

Conceptual Comprehension: They facilitate the visualization and explanation of variable behaviours, helping in the conceptual cognition of physical phenomena.

Preliminary Assessment: Sketched graphs are valuable in initial data scrutiny, helping to identify potential patterns and correlations before going into comprehensive quantitative analysis.

Educational Applications: Sketched graphs find frequent use in educational contexts to introduce principles and concepts without the complexity of precise numerical data.

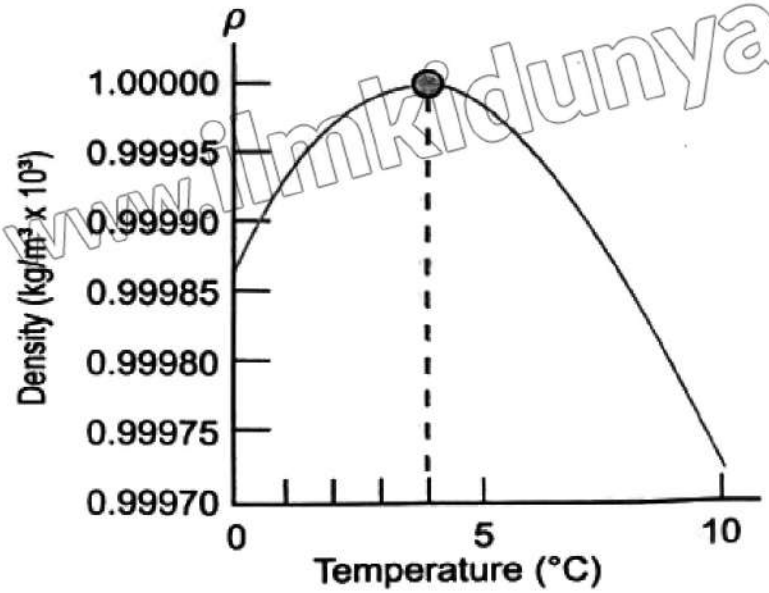
Instant Visualizations: Generating sketched graphs offers a swift and effective means of illustrating relationships, making them beneficial for brainstorming, presentations, or rapid analyses.

16.5.2 Drawn Graphs with Labelled and Scaled Axes

Drawn graphs are quantitative, graphical representation of the relationship between two variables. They have labelled and scaled axes. They give very important quantitative information about the quantities being plotted. The scale tells the intervals of quantities on each axis using a suitable unit of measurement. Scaled axes ascertain that the data points are correctly represented with relation to each other. The scale determines the gaps at which values are placed along each axis. This ensures the precise measurements andcomparisons between different data points. For example, the temp-density graph's scale on the x-axis shows temperature, each mark denotes increments of one-degree celsius, ensuring that temperature gaps are accurately reflected.

The density of water at different temperature

Temperature (°C)	Density (gcm^{-3})
100	0.9584
50	0.9880
25	0.9970
10	0.9997
6	0.9999
4	1.0000
0 (water)	0.9998
0 (ice)	0.9167



Graph 16.6: Scaled graph

It is more convenient to read patterns of increasing or decreasing order from a scaled graph, than by using a tabulated data.

Drawn graphs are useful for representing quantitative measurements, where numerical data is plotted. This includes variables such as density, distance, mass, temperature, or any other measurable quantity. By plotting data points on a graph, patterns, trends, and anomalies become visually apparent, aiding in data interpretation.

Drawn graphs are effective means for communicating results to a wide range of audience like general public. They successfully convey complex information in a visually accessible format, enabling easier understanding and interpretation of data.

16.5.3 How to calculate slope of a graph

1. Pick any two points on the line: Label them as Point 1 (x_1,y_1) and Point 2 (x_2,y_2).
2. Find the vertical change (rise): Subtract the y-values: y_2-y_1.
3. Find the horizontal change (run): Subtract the x-values: x_2-x_1
4. Divide rise by run to get the slope:

$$\text{Slope} = tan\theta = \frac{\text{perpendicular}}{\text{base}}$$

$$m = tan\theta = \frac{y_2 - y_1}{x_2 - x_1}$$

This slope tells you how steep the line is. Positive means uphill, negative means downhill, zero means flat, and undefined means vertical.

KEY POINTS

- Analytical chemistry is mainly related with the chemical characterization of materials.
- Quantitative analysis deals with the determination of how much of one or more constituents is present.
- There are two types of errors in data collection and in analysis leading to faulty results namely Random Errors, and Systematic errors.
- Random errors occur due to natural variability in the measurement process and can be caused by factors such as limitations in the measuring instrument, fluctuations in environmental conditions etc.
- Systematic error can be caused by an imperfection in the equipment being used.
- Error that can occur due to negligence or carelessness during the measurement process.
- Systematic errors cannot be eliminated by repeat trials, because systematic errors are consistent.
- There are two types of graphs; sketch graphs and drawn graphs.
- Drawn graphs are useful for representing quantitative measurements, where numerical data is plotted.

EXERCISE

1. Multiple Choice Questions (MCQs)

i. Which of the following is the example of qualitative analysis?
   a) Volumetry
   b) Spectrophotometry
   c) Gravimetry
   d) Precipitation

ii. During titration the burette is not correctly calibrated. The error that will generate due to this is categorised as:
   a) Random error
   b) Systematic error
   c) Indeterminate error
   d) Rough error

iii. Which error can be reduced or eliminated by taking repeat trials?
   a) Random errors
   b) Systematic error
   c) Determinate error
   d) Rough error

iv. Which one of the following gives more information?
   a) Information on x-axis
   b) Information on y-axis
   c) Information in un-scaled graph
   d) Information in scales graph

v. The density of water ___ from 4 °C to 0 °C.
   a) Increases
   b) Decreases
   c) Remains constant
   d) Fluctuates irregularly

2. Short Answer Questions

i. Explain qualitative and quantitative analysis with examples.

ii. What are random errors, write few examples.

iii. How does the error propegates? Give an example.

iv. Explain with one example each of dependent variable and independent variable.

3. Long Answer Questions

i. Explain errors and their types in detail.

ii. What are few methods to eliminate errors in analysis.