MODULE 1 FUNDAMENTALS OF PRE ANALYSES

UNIT 1 THEORY OF ERRORS

CONTENTS

1.0 INTRODUCTION

Analytical chemistry is a specialised aspect of chemistry that deals with both qualitative analysis (investigating what are the constituent of a given sample) and quantitative analysis (investigation on how much of constituents are in the sample). When reporting the quantitative measurement of an experiment, there are normally some controllable and uncontrollable variations that accompany such measurement. These forms of variations are termed **Errors**. This unit examines the concept of errors quantitatively in analytical chemistry.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- define what an error is
- state various types of errors
- express those errors correctly when reporting and
- report quantitative values to the correct number of significant figures.

3.0 HOW TO STUDY THE UNIT

1. You are expected to read carefully through this unit at least twice before attempting to answer the self-assessment questions or tutor- marked assignment.
2. Do not look at the solution given at the end of the unit until you are satisfied that you have done your best to get all answers.
3. Share your difficulties with your course mates, facilitators and by consulting other related material, particularly the internet.
4. Note that if you follow the instructions you will feel self fulfilled that you have achieved the aim of studying this unit. This should stimulate you to do more.

NOTE: All answers to activities and assignment are at the end of this book. This applies to every other unit in this book

4.0 WORD STUDY

Errors: are variations observed when the experimental value differs from the true value.

Random error: are due to the limitations of physical measurement and cannot be avoided.

Systematic error: are errors that can be determined and corrected.

5.0 MAIN CONTENT

5.1 Definition of an Error

Error is a controllable and uncontrollable variation observed when comparing the measured values to the true value. The error affects the accuracy and precision of a measured quantity.

5.2 Types of Errors

There are basically two types of errors: systematic errors and random errors.

5.2.1 Systematic Errors

These are otherwise known as determinate errors because they can be determined and corrected.

Example 1 Using pH meter that has been incorrectly standardised. If the pH buffer used is 14.07 but mistaken for 14.00, therefore a medium measured as 12.48 is actually 12.41 and a value put as 13.74 is actually 13.67. Any measurement with such pH meter must be buffered with a factor of 0.07.

5.2.1.1 Commonly Encountered Systematic Error

Systemic Errors commonly encountered in the course of laboratory exercises include:

(a) **Instrumental Error**
This occurs when faulty equipment and weight as well as glassware used are not calibrated or wrongly calibrated.

(b) **Operative Errors**
These are errors traced to the operators (personnel error), either as a result of the inexperience of the personnel involved or the operator not been careful enough. It may be mathematical error in the calculation or prejudice in estimating measurement.

(c) **Methodic Errors**
These are errors inherent in the analytical method or procedure used. These pose serious problems for an analyst. They include errors such as co-precipitation with impurities, incomplete reaction, impurities in the reagents used, etc. Methodical errors are also correctable.

5.2.1.2 Prediction and Correction of Systematic Error

There are various ways to detect and correct systematic error:

(a) Analyse samples of known composition.
Your method should reproduce the known answer; if not there is a problem with the method or equipment used.

(b) Analyse "blank" Samples.
Samples containing none of the parameter been sought. If you observe a non zero result, it means your method is responsible for more than what you intended.

(c) Use different analytical methods to measure the same quantity. If the results do not tally, it means there is an error associated with one of the methods.

(d) Let different operators of varying capabilities in different laboratories (using the same method or different methods) to carry out the same analysis. Disagreement or variations of high magnitude indicate error traceable to operators or equipment used.

5.2.2 Random Error

These errors are also known as indeterminate errors. They are errors due to the limitations of physical measurement and cannot be avoided. A better experiment or replicated experiment may reduce the magnitude of these types of errors, but cannot eliminate it totally. These types of errors are also called accidental errors. The errors are indicated by small differences in successive measurements made by the same analyst under almost identical experimental conditions. Random errors cannot be predicted or estimated. They can be either positive or negative.

Examples of Random Errors

(a) The type of variation associated with the same analyst reading the same absorbance scales many times

(b) Variation associated with the three or four different analyst reading the same measuring scale or reading the lower measurement of a volumetric flask. Obviously they would report varying values reflecting subjective interpolations between markings.

It has been observed that these types of errors always follow random distribution; hence mathematical laws of probability can help in arriving at conclusions regarding the most probable results in a series of measurements.

On a general note, errors affect the precision and accuracy of a measured quantity thereby raising questions on the integrity of the reported values.

5.2.3 Expressing Accuracy of a Measurement

There are various ways by which accuracy of a measurement can be expressed, these include:

(a) Absolute Error or Absolute Uncertainty

This is variation or difference shown between the true value and the measured value. It is reported in the same units as the measurement.

Example If a 4.97mg of an analyte is analysed as 4.91mg, the absolute error is 0.06

It becomes mean error if the measured value is the average of several measurements.

(b) Relative Error or Relative Uncertainty

It is an expression comparing the absolute uncertainty to the size of its associated measurement or absolute error expressed as percentage of the true value.

Example From (a) above, the relative error in the analysis is

$$\frac{4.91}{4.97} \times \frac{100\%}{1} = 98.79\%$$

The relative accuracy can then be deduced as follow.

$$\frac{4.91}{4.97} \times \frac{100\%}{1} = 98.79\%$$

Note that

i. the relative accuracy and relative error always give 100% if summed together

ii. relative errors can be expressed (as shown above) as parts per hundred (i.e in%) or parts per thousand (ppt).

Example If the result of an analysis, is 29.74μg and is compared to the true value of 30.15 μg. Calculate the relative error in part per hundred and part per thousand.

Solution

Absolute Error = 29.74 - 30.15 = -0.41

$$\text{Relative Error in pph} = \frac{0.41}{30.15} \times 100 = 1.36\%$$

$$\text{Relative error in ppt} = \frac{0.41}{30.15} \times 1000 = 0.45 \text{ ppt}$$

6.0 ACTIVITY

i. Define the following terms
   (a) Error (b) Blank sample (c) Part per thousands.
ii. Briefly describe the major types of error known.
iii. Differentiate between qualitative and quantitative analysis.

7.0 SUMMARY

In this unit, you have learnt that:

i. Analytical chemistry deals with both quantitative and qualitative analysis.
ii. Errors are variations that naturally accompanied the experiment performed.
iii. Two major types of errors are known namely: determinate and indeterminate errors.
iv. Determinate errors can be predicted and accounted for, as well as corrected while indeterminate errors are accidental and inherent in the experiment itself and hence cannot be corrected.
v. Absolute and relative errors or uncertainty are two ways of expressing the accuracy of the measured values.
vi. Relative error can be expressed in terms of part per hundred (pph), that is % or parts per thousand (ppt).

8.0 ASSIGNMENT

1. Differentiate between quantitative and qualitative analysis
2a. Explain the meaning of the following terms
    (i) Analyte (ii) Blank (iii) operational error.
b. Differentiate between random and systematic errors.
3. Briefly explain ways by which systematic error can be predicted and corrected.

4. A standard serum containing 400 mg/L of chloride was analysed 398, 397, 396 and 399mg/L were obtained. Calculate (a) the mean value (b) absolute error and relative error in percent.

9.0 REFERENCES/FURTHER READING

Christian, G. D. (1980). *Analytical Chemistry*. (3rd ed). New York: John Wiley and son.

Harris, D.C. (1995). *Quantitative Chemical Analysis*. (4th ed). New York: Freeman and Company.

Khan, I. A. & Khanum K. (1994). *Fundamentals of Biostatistics.* Nagar: Ukaaz Publications.

Nwachukwu, V. O. (2006). *Principle of Statistical Inference*. Port Harcourt: Peace Publishers.

UNIT 2 STATISTICAL TREATMENT OF DATA

CONTENTS

1.0 INTRODUCTION

In the past, great chemist had faced various challenges, which include the necessity to analyse a reasonable large number of samples in so many monitoring effort so as to ensure representative coverage, selection of appropriate method out of many known suitable analytical techniques, the problem of variations in reported values from different methodologies employed in analysing the same samples, coping with different measured values reported by the same operation in replicate and unthinkable interrelationship that could exist among the data generated. These problems and many others had forced scientist to just collecting “base line” data for referral use, against which the future can be assessed.

However, it is now realised that with the statistical analytical tools, various problems could be solved. Statistics enables analytical chemist to accept conclusions that have high probability of being correct and to reject conclusions that are doubtful. Hence, statistical treatment of data helps in ascertaining the significance and integrity of values reported.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- list appropriate statistical tools available for data handling
- define various statistical terms and state their importance and
- use various statistical tools in interpreting data and arrive at a safe conclusion.

3.0 HOW TO STUDY THE UNIT

1. You are expected to read carefully through this unit at least twice before attempting to answer the self-assessment questions or tutor- marked assignment.
2. Do not look at the solution given at the end of the unit until you are satisfied that you have done your best to get all answers.
3. Share your difficulties with your course mates, facilitators and by consulting other related material, particularly the internet.
4. Note that if you follow the instructions you will feel self fulfilled that you have achieved the aim of studying this unit. This should stimulate you to do more.

NOTE: All answers to activities and assignment are at the end of this book.

4.0 WORD STUDY

Accuracy:	is the difference between true and actual value of an experiment.
Precision:	is the degree of agreement between replicate measurements of the same quantity.
Variance:	is the difference between expected and actual values.
Significant figure:	is the minimum number of digit required to express a given value in a scientific way with a measured precision

5.0 MAIN CONTENT

5.1 Definition of Various Statistical Terms

5.1.1 Significant Figures

It can be defined as the minimum number of digit required to express a given value in a scientific way with a measured precision.

This concept is very important in conveying the actual meaning and status of each digit. However, many individual are not properly schooled in the use of significant figures and hence make figures in experimental reports appear confused.

The digit zero [0] can be a significant part of a measurement depending on where it occurs.

Generally, zeros are significant if:

i. they occur in the middle of a number
ii. they occur at the end of a number on the right-hand side of a decimal point.

Note that the number of significant figures in a measurement is independent of the placement of decimal point.

Example: The significant zeros are underlined. 7**0**4, 0.07**0**4, 0.7**0**4, 7**0**4**0**, 0.7**0**4**0**. Ambiguity arises when a figure like 92500 is written in respect of the significant figures. However, it could be written in any of the following ways.

9.25×10^4 - 3 significant figures
9.250×10^4 - 4 significant figures
9.2500×10^4 - 5 significant figures

Note that the first uncertain figure is the last significant figure.

5.1.2 Rounding Off

The problem often encountered in significant figures is when an arithmetical operation takes place and when the answer is to be rounded off. The operation is either addition/subtraction, multiplication/division.

The general rule is: Rounding should only be done on the final answer (not intermediate results) to avoid build-up of round-off errors.

5.1.3 Addition and Subtraction

Rule: Express all numbers with the same exponent and align all numbers with respect to the decimal point.

Round-off the answer according to the number of decimal places in the number with the fewest decimal places.

Example (a)

$$\begin{array}{r} 14.344137 \\ +\,17.347799 \\ +\,44.313 \\ \hline 76.50\underbrace{\;4936}_{\text{not significant}} \\ \hline \end{array}$$

The final answer is 76.505

When adding or subtracting numbers expressed in scientific notation, all numbers should first be converted to the same exponent.

Example (b)

$$\begin{array}{r} 1.373 \times 10^{5} \\ +\,5.314 \times 10^{3} \\ +\,0.798 \times 10^{6} \end{array}$$

Convert to the same exponent

$$\begin{array}{lcl} 1.373 \times 10^{5} & & 1.373 \times 10^{5} \\ 5.314 \times 10^{3} & \Longrightarrow & 0.05314 \times 10^{5} \\ 0.798 \times 10^{6} & & \underline{7.98 \times 10^{5}} \\ & & \underline{9.40614 \times 10^{5}} \end{array}$$

Round off to fewest decimal point.

So the final answer is 9.41×10^{5}.

5.1.4 Multiplication and Division

The operation should be limited to the number digit contained in the number with the fewest significant figures.

Example (a)

$$\begin{array}{r} 3.26 \times 10^{-5} \\ \times\ 1.78 \\ \hline 5.80 \times 10^{-5} \\ \hline \end{array}$$

(b)

$$\begin{array}{r} 4.3179 \times 10^{12} \\ \times\ 3.6 \times 10^{-19} \\ \hline 1.6 \times 10^{-6} \\ \hline \end{array}$$

(c)

$$\begin{array}{r} 34.60 \\ \times\ 2.46287 \\ \hline 85.2153020 \\ \hline \end{array}$$

Note that the power of 10 has no influence on the number of figures that should be retained.

5.1.5 Rounding Off Rule

Note that the rounding off should be done on the final answer, when the arithmetic operation must have been carried out.

The following rules validate rounding off operation.

1. If the digit following the last significant figure is greater than 5, the number is rounded up to the next higher digit.
2. If the number is less than 5, the number is rounded to the present value of the last significant figure.

Example: 9.47 = 9.5 in two significant figures.
9.43 = 9.4

3. If the last digit is a 5, the number is rounded off to the nearest even digit.

Example: 4.65 = 4.6 not 4.7
4.75 = 4.8
4.55 = 4.6

5.2 Ways of Expressing Precision

Statistics has enabled scientist to accept or reject conclusions on figures depending on the degree of precision carried or attached by the numerical report.

Precision is defined as the degree of agreement between replicate measurements of the same quantity. There are various tools that can be use in expressing the precision. These include: average deviation, variance, standard deviation etc.

5.2.1 Average Deviation (A.D)

It is one of the methods of showing dispersion or way of ascertaining the deviation from the central values. It is otherwise called Mean Deviation. It helps further in measuring distribution that is based upon all the items in a distribution.

Mean deviation or average deviation =

$$\text{AD} = \frac{\sum |x - \bar{x}|}{n} \quad \text{or} \quad \frac{\sum |d\bar{x}|}{n}$$

$d\bar{x}$ = deviation from mean
n = number of observation
x = observation
$\bar{x}$ = sample mean

$$\text{Coefficient of mean deviation} = \frac{Mean\ Deviation}{Mean}$$

As with accuracy, precision measurement such as average deviation can be expressed as an absolute figure or as a relative figure (%, pph, ppt etc).

5.2.2 Variance

This is simply called Mean Square Deviation. Variance is an important measure in the quantitative analysis of data. It helps in isolating the effect of various factors. It also helps in developing some statistical theories.

$$Variance\left(S^2\right)=\frac{\sum(x-\bar{x})^2}{n-1} \quad or \quad \frac{\sum(d\bar{x})^2}{n-1}$$

$\bar{x}$ = arithmetic mean
n = number of observation

5.2.3 Standard Deviation (SD)

It is the most commonly used absolute measure of dispersion. It measures how closely the data clustered about the mean.

Note: The smaller the standard deviation the more closely the data are clustered about the mean i.e. homogeneity is observed when standard deviation is small.

Therefore, standard deviation measures the spread in a set of observation. Standard Deviation (S) is simply the square root of the variance.

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{\Sigma(d\bar{x})^2}{n-1}}$$

Coefficient of Variation: The standard deviation is an absolute measure of dispersion. It is expressed in terms of unit in which the original data is collected. For instance the standard deviation of length of fish is different from standard deviation of weight of fish. To enable comparison of the two, there is the need for conversion into relative measure. This relative measure of dispersion is known as coefficient of variation (C.V).

$$C.V=\frac{S}{\bar{x}}\times 100$$

Where S = Standard deviation
$\bar{x}$ = Mean

5.2.4 Standard Deviation of Mean

This is otherwise known as standard error of mean (SEM).

$$SEM=\frac{S}{\sqrt{N}}$$

Where S = Standard deviation
N = number of observation.

Note that when the sample given during the measurement of dispersion is less than 10, we use n – 1 but if it is more than10, we use n.

5.2.5 Student's T - Test

This is a statistical tool used most frequently to compare the mean values from experimental procedure. It also helps in expressing confidence interval. This is the range within which the true value might fall within a given probability. The limit of this range is called confidence limit. The likelihood that the true value falls within the range is called the probability or confidence level, usually expressed as a percentage.

A statistical t-value is calculated (t_{cal}) and compared with the tabulated t-value (t_{tab}). If the calculated t-value exceeds the tabulated t- value, then there is a significant difference between the results of the two methods at that confidence level. If it does not exceed the tabulated t-value, then we can predict that there is no significant difference between the methods.

There are three ways by which t-test can be used:

(a) T-test when a standard or true value is known

$$\pm t = (\bar{x} - \mu)\frac{\sqrt{N}}{S}$$

Where $\bar{x}$ = mean value
μ = true value
N = number of observations
S = standard deviation

(b) T-test when comparing replicate measurements

$$\pm t = \frac{\bar{x}_1 - \bar{x}_2}{Sp}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$Sp = \sqrt{\frac{\Sigma(x_{li} - \bar{x}_l)^2 + \Sigma\left(x_{i_2} - \bar{x}_2\right)^2 + \cdots \Sigma(xi_k - x_k)^2}{n_1 + n_2 - K}}$$

Where, $\bar{x}_1$, $\bar{x}_2$,.............$\bar{x}_n$, are mean values of each of the set.
Sp = Pooled standard deviation
X_{li}, $xi_{2,}$... ,xi_k = Individual value in each set.

K = sets of analyses.

(c) T-test when comparing individual difference.

This case applies when we use two different methods to make single measurement on several different samples. No measurement has been duplicated.

$$t = \frac{\bar{d}}{Sd}\sqrt{n}$$

$$Sd = \sqrt{\frac{\sum (di - \bar{d})^2}{n-1}}$$

d_1 = The individual difference between two methods for each samples with regards to sign.

$\bar{d}$ = The mean of all the individual difference.

5.2.6 F – Test

This is a test designed to investigate whether there is a significant difference between two methods based on their standard deviation. It is defined in terms of variance.

$$F_1 = \frac{S_1^2}{S_2^2}$$

Where $S_1^2 > S_2^2$. If the calculated F value exceeds tabulated F value at a given confidence level, there is a significant difference between the variance.

5.2.7 Correlation

This is a statistical tool with the help of which the relationship between two variables is studied. Indeed, correlation studies also help to show degree of any association (quantitatively) between two sets of variables.

With this knowledge, one can predict if the existence of trend in one variable will affect the other. There are three major ways by which correlation is carried out.

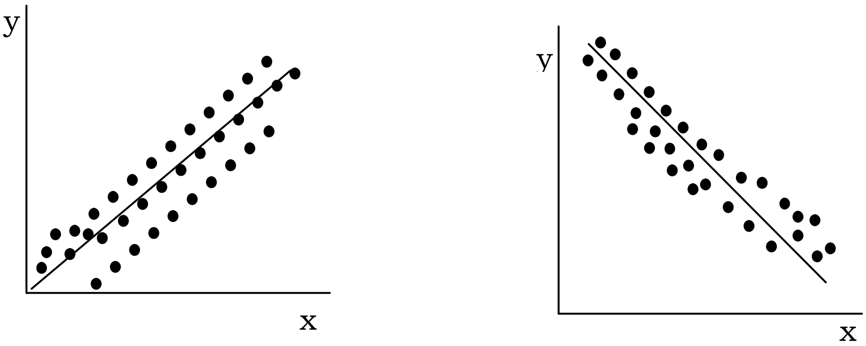
i. Scatter diagram method
ii. Graphic method
iii. Coefficient of correlation.

(i) Scatter Diagram Method

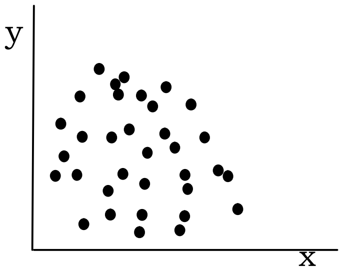
It is the simplest method for ascertaining correlation between two variables by plotting the values on a chart known as scatter diagram.

In plotting this diagram, note that the X variable is the independent variable while on y-axis you plot dependable variable.

Example: The height of plant is on x axis while the number of flower is on y-axis. The following types of scatter graph are commonly obtained.



(a) Positive Correlation **(b) Negative Correlation**

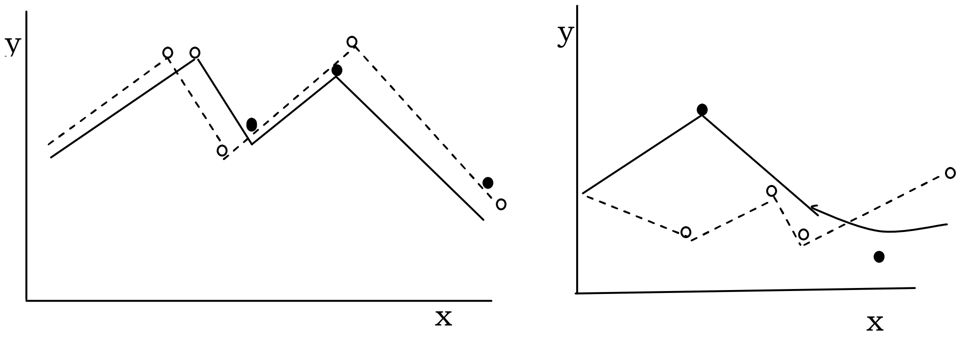


(c) No Correlation

Figure 1.0 Diagram of scatter patterns of correlation.

(ii) Graphic Method

This is a mutual mathematical graph in which graphs of two variables in question are plotted and are related to one another. The graphs would either show positive or negative correlation.



(a) Positive Correlation (b) Negative Correlation

Figure 2.0 Graphical Pattern of Correlation

(iii) Coefficient of Correlation

This is when the degree of relationship can be established by calculating a coefficient which always gives a quantitative measure of the degree of closeness between two variables. This postulate is the basis for ranking numerical measure of degree of correlation. One of such numerical measured is Pearson Correlation coefficient, r.

$$r = \frac{\sum (x-\bar{x})\ (y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \ \sum y-\bar{y})^2}}$$

x = the independent variable
y = dependent variables
r = correlation coefficient
$\bar{x}$ and $\bar{y}$ are the mean values of independent and dependent variables, respectively.

However for simplification purposes, the r can be re-written as:

$$r \quad \frac{\sum xy - \frac{\sum x \sum y}{n}}{\left[\sqrt{\sum n^2} - \frac{(\sum x)^2}{n}\right]\left[\sum y^2 - \frac{(\sum y)^2}{n}\right]}$$

The r calculated would show the degree of the inter-relationship as indicated in Table 1 below.

Table 1.0 Interpretation of Degree of Correlation

Degree of Correlation	Positive	Negative
Perfect Correlation	+1	-1
Very high Correlation	+ 0.9 or more	-0.9 or more
Sufficient Correlation	+0.75 to 0.9	-0.75 to - 0.9
Moderate Correlation	+0.6 to 0.75	-0.6 to -0.75
Possible Correlation	+0.3 to 0.6	-0.3 to -0.6
Possibly No Correlation	Less than 0.3	Less than -0.3
Absence of Correlation	0	0

6.0 ACTIVITY

i. Briefly describe various ways by which precision can be expressed.

ii. Explain the meaning and importance of the following terms:

a. Precision and Accuracy
b. Standard Error Mean
c. Significant figures

7.0 SUMMARY

In this unit, you have learnt that:

A. Statistical handling of data is very important.
B. Significant figures help to reveal the status of all digits in the measurement as well as indicating the degree of its uncertainty.
C. Various ways by which precision in a figure can be expressed include average deviation, standard deviation, variance, etc.
D. Both T and F tests are capable of revealing the significance of the difference between mean values from different experiment.
E. Various ways of showing correlation between two variables and their merits are possible.

8.0 ASSIGNMENTS

i.a. Differentiate between F and T test
b. Calculate the average deviation and the average relative deviation of the following set of analytical results: 18.40g, 18.37g, 18.43g and 18.39g.
ii. How many significant figures does each of the following number have?
 (a) 2000.06 (b) 6.030×10^{-4} (c) 7.30×10^{10}
iii. Distinguish between accuracy and precision
iv. Replicate samples of silver alloys are analysed and determined to contain 95.67, 95.61, 95.71 and 95.60% of silver metal. Calculate (a) average deviation in ppt, (b) the standard deviation and (c) the relative standard deviation of mean

9.0 REFERENCES/FURTHER READING

Christian, G. D. (1980). *Analytical Chemistry*. (3rd ed), New York: John Wiley and Son.

Harris, D. C. (1995). *Quantitative Chemical Analysis*. (4th ed). New York: Freeman and Company.

Khan, I. A. and Khanum K. (1994). *Fundamentals of Biostatistics*. Nagar: Ukaaz Publications.

Nwachukwu, V.O. (2006). *Principle of Statistical Inference*. Port- Harcourt: Peace Publishers.

UNIT 3 THEORY OF SAMPLING

CONTENTS

1.0 INTRODUCTION

There are certain important steps that naturally precede chemical analysis if the result or data generated is to have any significance. These steps include:

(i) Sampling
(ii) Production of a homogenous mixture for analysis and
(iii) Drying the collected sample.

Sampling wrongly done will not yield any meaningful result no matter how painstaking or laborious the analysis is. Materials to be analysed more often than not, exist not only in a large size but also in complex non-homogeneous forms. Hence, attention, commitment and expertise are required to obtain a true representative sample. The number or size of the sample is all geared towards a desired result, which is achievable only when the problem of analytical exercise is well defined.

This unit covers the scope and principle employed in getting true sample, preparation of sample and Statistics of sampling.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- explain the basic principles of sampling
- state types of sampling and
- list and explain the various methods of sampling

3.0 HOW TO STUDY THE UNIT

1. You are expected to read carefully through this unit at least twice before attempting to answer the self-assessment questions or tutor- marked assignment.
2. Do not look at the solution given at the end of the unit until you are satisfied that you have done your best to get all answers.
3. Share your difficulties with your course mates, facilitators and by consulting other related material, particularly the internet.
4. Note that if you follow the instructions you will feel self fulfilled that you have achieved the aim of studying this unit. This should stimulate you to do more.

NOTE: All answers to activities and assignment are at the end of this book.

4.0 WORD STUDY

Sampling: are the operations involved in procuring a laboratory size that is a true representative of “a whole lot” for a particular analytical exercise.

Random sampling: is a method of sampling in which each item of the population has equal chance of being included in the samples.

Stratified Sampling: It is a technique employed when population is heterogeneous with respect to the variables under study.

5.0 MAIN CONTENT

5.1 Basic Principles of Sampling

Sampling can be described as the operations involved in procuring a laboratory size that is a true representative of “a whole lot” for a particular analytical exercise.

Sampling is indeed the most difficult step in the entire analytical process; however, it remains the only key to the success of the whole analytical programme.

Good sample irrespective of the type or method of sampling should possess the following properties:

i. A good sample must have the same characteristic or features with that of the original population from where it is selected.

ii. The nature of the sample must be the same with that of the population and must remain so throughout the analytical exercise.
iii. The number of samples should be large enough to make the result reliable.

Indeed, the method used in collecting a true representative sample depends on various factors which include:

a. The knowledge and experience of the analyst.
b. The result of survey on the nature, size and configuration of the site of materials to be sampled.
c. The level of sensitivity of the desired result.

It is important to note that, there is virtually no single technique that can satisfy all requirements in any sampling case. Modifications and combination of some techniques may be necessary in some sampling cases.

5.2 Sampling Techniques

Basically there are two main sampling techniques.

i. **Random Sampling**:This is a method of sampling in which each item of the population has equal chance of being included in the samples. Random sampling can be from either finite samples or infinite samples.
ii. **Stratified Sampling**: It is a technique employed when population is heterogeneous with respect to the variables (parameters) under study. The population is divided into varying homogeneous groups or strata and random sample is drawn from each stratum and pooled together.

5.3 Types of Samples

The following are various types of sample that can be collected.

i. **Grab Samples**: These are samples collected at specific spot at a site over a short period of time. These types of samples represent a “snap shot” in both space and time of the sampling area.

Grab Sampling can be (a) discrete grab samples i.e. samples taken at a selected location depth and time, (b) depth-integrated sample which are collected over a predetermined part or to entire depth of an area with respect to location and time.

Grab Sample is suitable either when the source is known to vary with time or when source composition varies in space.

ii. **Composite Samples**: These types of sample provide more representative sampling of heterogeneous matrices in which the composition of the analyte of interest may vary over a period of time and or space. Composite samples can be:

(a) Sequential (time): composites are collected using continuous and constant equipment (like pumping machine or by mixing equal volume of water collected at a regular time interval.

(b) Flow-Proportional: composite samples are collected at a rate proportional to flow rate.

iii. **Integrated Samples**: At times, information or result desired may be best provided by analysing mixture of grab samples collected from different points simultaneously or as nearly as possible using discharge- weighted methods such as Equal Width Increment (EWI) or Equal Discloses Increment (EDI).

Generally, sampling can be done manually or with the aid of instrument depending on various factors such as cost of the analysis, size of the sampling site and the numbers of samples to be procured.

5.4 Gross and Laboratory Size Samples

Irrespective of the state of materials to be sampled, the level of heterogeneity, the size, nature and volume of the sample and configuration of the site, the interest of the analyst is to generate gross sample from which laboratory size sample is obtained for analysis.

5.4.1 Gross Sample

Ideally, gross sample is a miniature replica of the bulk of materials to be analysed. It corresponds to the "whole lot" both in chemical composition and in particle size distribution. A certain portion of the whole must be removed through any of the sampling methods earlier discussed. The sample may be grab or composite depending on the judgement of the analyst. The competence and expertise required for obtaining the gross sample vary and depend on the situation at hand. This range from sampling homogeneous situation of liquid and gases, and sampling particulate solids, to sampling of metal and alloys.

The size of the gross sample needs not to be larger than necessary. The size is determined by the following factors:

i. the uncertainty that can be tolerated between the composition of the samples as a whole
ii. the degree of heterogeneity of materials being sampled and
iii. the level of particle size at which heterogeneity begins.

5.4.2 Laboratory Size Sample

This is the ultimate sample on which analysis is carried out from non- homogeneous material. The gross sample may weigh several hundreds of kg or more. A laboratory size that is almost one over thousands or less is obtained. Diminutions in particles size is essential as the weight of the sample is decreased to ensure that the sample composition continue to be the representative of the original materials.

The sample obtained upon arrival at the laboratory received further treatment before it is eventually analysed. The integrity of sample to maintain chain-of-custody procedure must be ensured.

5.4.2.1 Coning and Quartering Method of Sample Selection

It is a method of sampling selection which aims at reducing the samples without creating a systematic bias. The technique involves pouring the sample so that it takes on a conical shape, and then flattening it out into a cake. The cake is then divided into quarters and two quarters which face opposite one another are discarded, whilst the other two are combined and constitute the reduced sample. The same process is continued until a reasonable amount of material is obtained. Analyses are made with respect to the sample obtained.

The chain of custody procedure is the following:

1. Sample label (including bar code label)
2. Sample seals
3. Fields log book
4. Chain-of-custody books
5. Sample analysis request sheet
6. Sample delivery to the laboratory
7. Receipt and logging of sample
8. Assignment of sample for analysis
9. Disposal

The chain-of-custody procedure helps in identifying the source of contamination (if any) and help the analyst in planning.

5.5 Treatment or Preparation of Sample

Samples brought to the laboratory require further treatment before analysis commences. This is due to the need to convert the sample from the nature in which it exist at the site of sampling to the form in which it can be analysed. The treatments also help in eliminating the possible sources of contamination and sample degradation that could lead to sample destruction. The treatment of samples also ensures the homogeneity of samples, so that any small portion removed for the analysis will be identical to any other fraction.

Major activities during the preparation of samples include.

a. Concentration: This implies reducing the water content of a sample material (drying)
b. Dissolution: This involves converting sample material in solid form to solution
c. Grinding and Crushing: This involves reducing the size of particles of sample materials. It helps increasing the surface area thereby allowing effective attack of reagents during reaction.
d. Mixing solids laboratory samples: It is essential that solid materials be thoroughly mixed in order to ensure random distribution of the components in the analytical sample.

The figure below shows sample treatment chart

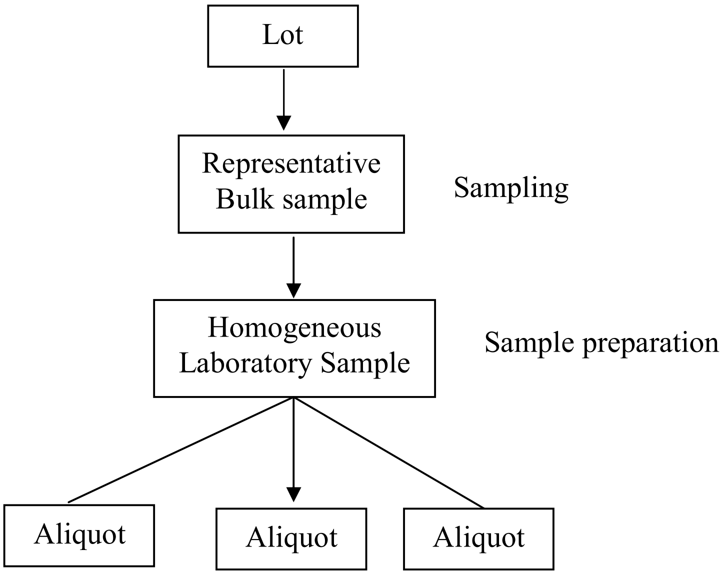


Fig. 1.0: Sample treatment flow chart

5.6 Statistics of Sampling

The use of statistics has gone a long way in helping to unravel some of what could be a difficult decision to make at the sampling site. Such decisions as what should be the sample size or what number of samples to collect could be statistically handled.

(i) **What should be the size of sample?**

When n particles are drawn from mixture of two kinds (such as liver tissue particle and droplet of water), the sampling standard deviation will be

$$\sigma_n = \sqrt{npq}$$

Where p and q are the fractions of each kind of particles present

The relative standard deviation is $\sigma_n = \sqrt{pq/n}$

The relative variance $(\sigma_n/n)^2$ is therefore

$$R^2\left[\frac{\sigma_n}{n}\right] = \frac{PE}{n} = nR^2 = pq$$

Where mass of the sample drawn (m) is proportional to the number of the particles drawn

$mR^2 = K_S$

K_S = Sampling Constant

$$m = \frac{Ks}{R^2}$$

(ii) **What number of samples to collect?**

The sampling contribution to the overall uncertainty can be minimised by analysing more samples

Rearranging student's t-test will allow us to know the number of sample required to meet a desired confidence levels

$$\mu - \bar{x} = \frac{ts_s}{\sqrt{n}}$$

$$n = \frac{t^2 s^2}{(\mu = x)^2}$$

$Let\ \mu \quad \bar{x}$ be = e

$$n = \frac{t^2 s^2}{e^2}$$

where μ = true population mean

$\bar{x}$ = measured mean

n = number of samples needed

S^2 = variance

5.7 Sample Decomposition

This is converting the accessible sample matrices to the form which is accessible to instrument. There are various methods employed in decomposition. This varies depending on the chemical nature of the sample.

5.7.1 Inorganic Solids

(a) Strong mineral acids are good solvent for many organics. Acids such as hydrochloric acid, nitric acid, hydrofluoric acid, perchloric and sulphuric acids are commonly used.
(b) Fusion of Inorganic Materials with acid or basic flux in a molten state is the only way through which some materials are digested. The sample is mixed with flux in a ratio of about 1 to 10, or 20, sample to flux, the combination is heated in an appropriate crucible until the flux becomes molten, the cooled solid is then afterward dissolved in dilute acid or water.

5.7.2 Organic Materials

Organic material such as plant or animal tissue or biological fluid are usually decomposed by

i. Wet Digestion: This is achieved by boiling sample materials with oxidizing mineral acid or mixture of acids.
ii. Dry Ashing: This involves heating the materials at high temperature (400-700^{0}C) in a muffle furnace until it is turned to ashes. This is later dissolved in mineral acid or water.

6.0 ACTIVITY

i. What is the chain-of-custody procedure?
ii. Explain the meaning of the following terms (a) Coning (b) Quartering (c) Wet digestion**?**
ii. Write brief notes on (a) Random Sampling (b) Stratified Sampling (c) Grab Sampling (d) Composite Sampling

7.0 SUMMARY

In this unit, you have learnt that:

- The definition of basic principle of sampling
- The methods of sampling as well as the types of samples
- Detail of chain-of-custody proceeding
- Preparation of samples
- Decomposition of sample
- Statistics of sampling.

8.0 ASSIGNMENT

a. What are the two principal means of dissolving inorganic material?
b.
   i. Describe the principle of ashing and wet digestion.
   ii. What are the procedures of chain-of-custody?
   iii. Justify this statement: “sampling is the key to a successful analysis.
   iv. Explain in detail, sample preparation.

9.0 REFERENCES/FURTHER READING

Christian, G. D. (1980). *Analytical Chemistry*. (3rd ed), New York: John Wiley and Son.

Harris, D. C. (1995). *Quantitative Chemical Analysis*. (4th ed). New York: Freeman and Company.

Khan, I. A. and Khanum K. (1994). *Fundamentals of Biostatistics*. Nagar: Ukaaz Publications.

Nwachukwu, V.O. (2006).*Principle of Statistical Inference*. Port- Harcourt: Peace Publishers.