Dr. Sattar Rajab Majeed

(<u>Analytical Colculation</u>)

<u>MSc. Semester 2</u>

5 The Evaluation of Data and Results Reliable in Quantitative Chemical Analysis

Every physical measurement is subject to a degree of uncertainty which can be decreased to an acceptable level. In this stage of study, it is only required to recognise these uncertainties in the data that obtained after experimental effort.

From the practical view of the chemist, he attempts to repeat his measurement, several times to find out the extent of closeness of this data to each other. On other hand, if the chemist has obtained several identical measurements, he will encounter a problem of picking up the best value of his measurements. The variations among the results ought to provide a measure of reliability in the best value that is chosen.

5.1 Types of errors in analytical results.

The errors that associated with chemical analysis are classified into two main groups:

5.1.1 Determinate errors

These errors have a definite value which can be measured and accounted for. These errors are subclassified briefly into the following errors:

5.1.1.1 Personal errors:

These errors are the result of the ignorance, carelessness and physical limitations of the experimenter. They may arise from the use of an improper technique in transferring a sample, from disregarding of temperature corrections for a measuring item and from over or under

washing the precipitate. A physical limitation is found in the partially colour-blind person who has great difficulty in distinguishing colour changes during the process of the analysis. A personal error may be due to personal bias. There is a natural tendency for the experimenter to estimate or select the measurement or the result which serves his purposes. This type of error can be reduced by practice and by conscious use of objectivity in making observations.

5.1.1.2 Instrumental errors:

These errors are attributed to the imperfections of the items and equipments. Volumetric flasks, burettes and pipettes are frequently calibrated at a certain temperature to deliver or contain volumes exactly obtained at the calibrated temperature. Otherwise these items will be a source of determinate error.

5.1.1.3 Method errors:

These errors are often introduced from non-ideal chemical behaviour of the reagents and their reactions. These sources are the slowness of some reactions, incompleteness of others, instability of some species and the occurrence of side reactions which interfere with the measurement process. For example, failure of sufficient washing of the precipitate will be contaminated with foreign substances and have a high weight. On other hand, extra washing of the precipitate may cause loss in weight and give low results. A method error frequently encountered in volumetric analysis is due to the volume of the reagent which is consumed by an indicator to cause colour change that indicates the completion of the reaction.

Errors appear in a method are probably the most serious of determinate error since they are the most likely to remain undetected.

Method errors are reduced by:

i) **Analysis standard samples.**

ii) **Using more than one method in analysis and comparing results. For example iron can be determined gravimetrically and by volumetric titration method and comparing the obtained results.**

iii) **The use of different weights and different volumes for the procedure of analysis to get the actual results.**

iv) **Employing the blank to insure the interferents of the method. Blank is the system which contains all the species except the unknown (analyte) under the same conditions.**

5.1.1.4 Effect of determinate errors on analytical results:

Determinate errors generally fall into either constant or proportional errors.

Constant errors are independent of the size of measured quantity. These errors become more serious as the size of measured quantity decreases. These can be illustrated by the following example.

Suppose washing 500 mg of precipitate with 200 ml of water and the lost by washing is 0.50 mg. Therefore the relative loss by solubility

$$= -0.50 \times \frac{100}{500} = -0.1\%$$

While the relative loss of the same quantity(0.5 mg)for 50 mg of precipitate

$$= -0.50 \times \frac{100}{50} = -1\%$$

which is unacceptable value resulted from using little precipitate.

Another example of constant error is the amount of reagent required to bring about the colour change in volumetric analysis. This volume

remains the same regardless of the total volume of required reagent. Again, the relative error will be more serious as the total volume decreases. Therefore, constant errors are minimised by employing relatively large samples.

Proportional errors:

These errors are caused by the presence of interferents, for example: copper is analysed by reaction with potassium iodide. The presence of Fe^{3+} liberates also I_2 from potassium iodide:

$$2Cu^{2+} + 4I^{-} \longrightarrow 2CuI + I_2$$

$$Fe^{3+} + 2I^{-} \longrightarrow Fe^{2+} + I_2$$

The presence of Fe^{3+} will raise the percentage of Cu since the iodine produced will be a measure of the sum of copper and iron in the sample. If the sample size is doubled, the amount of iodine liberated by both the copper and iron will be doubled. While the absolute error will undergo a twofold increase, the relative error will remain unchanged.

5.1.2 Indeterminate errors (Random errors):

These errors arise from uncertainties in a measurements that are unknown and can not be controlled by the scientists. The effect of such uncertainties is to produce a scatter of results for replicate measurements. These errors therefore, are called random errors and follow the loss of probabilities. Therefore, some statistical bases and laws are used to evaluate the experimental results.

In the experimental process, a set of data is obtained which led to the results.

For instance, in the acid-base titration, a certain volume of acid or base is titrated with base or acid respectively, such as the titration of 10 ml of

Na_2CO_3 with HCl solution in the presence of methyl orange as indicator. The following data are obtained.

ml Na_2CO_3	**10**	**10**	**10**	**10**	**10**	**10**	**10**
ml HCl	**10.2**	**10.3**	**10.1**	**9.9**	**9.8**	**10.1**	**10**

Thus, different volumes of HCl are obtained in the titration of 10 of Na_2CO_3 solution. Which one of these volumes is reliable and what are the statistical bases and laws that can be employed to evaluate these measurements?

5.2 The Precision:

It means the agreement between numerical values of two or more measurements that have been done under the same conditions. It describes the reproducibility of the measurements and the extent of their closeness to each other. As the difference between the data is small, the high precision is obtained.

There are several statistical methods for expressing the precision for the obtained or attained measurements. The following methods are the simple expressions for the precision:

5.2.1 The mean, Arithmetic mean or average: ($\bar{X}$)

The mean, arithmetic mean and average are synonymous terms for the numerical value obtained by dividing the sum of a set of replicate measurements by the number of individual results in the set.

$$\bar{X} = \frac{\sum X_1 + X_2 + X_3 + \ldots\ldots\ldots X_n}{N}$$

where N= the number of individual measurements.

$X_1 + X_2 + X_3 + \ldots\ldots\ldots X_n$ are the individual data, and Σ is the summation term.

Ex: Calculate the mean for the titration of 10 ml of Na_2CO_3 with HCl solution:

0.1 N Na_2CO_3 solution (ml)	10.0	10.0	10.0	10.0	10.0	10.0	10.0
0.1 N HCl solution (ml)	10.2	10.3	10.1	9.9	9.8	10.1	10.0

The solution:

$$\bar{X} = \frac{\Sigma X_1 + X_2 + X_3 + \ldots\ldots\ldots X_n}{N} \qquad \bar{X} = \frac{10.2+10.3+10.1+9.9+9.8+10.1+10.0}{7} = 10.06$$

5.2.2 The median $(\bar{X}_m)$:

It is the central value of an odd number of a set of measurements or the average of the central pair for a set containing an even number of measurements.

Ex: Calculate the median of the two following sets of measurements:

7 13 16 <u>20</u> 23 27 29 :odd number

$\bar{X}_m = 20$

7 12 13 <u>16 18</u> 20 25 29 : even number

$$\bar{X}_m = \frac{16+18}{2} = 17$$

The median may equal to the mean, but they are always not identical especially when the number of measurements in the set is small.

5.2.3 The spread or range (w):

It is the numerical difference between the highest and the lowest values of measurements.

For the following data :10.06, 10.08, 10.10 and 10.20

The range (w) = 10.20-10.06 = 0.14

As long as the range is small, the measurements are more précised.

5.2.4 Deviation from mean (d):

It is a numerical difference (regardless of sign) between any experimental value and the mean for the set of data that includes the value.

$d = (X_i - \bar{X})$ X_i = the individual value .

5.2.5 Deviation from median (dm):

It is a numerical difference (regardless of sign) between any experimental value and the median for the set of data that includes the value.

$dm = (X_i - \bar{X}_m)$

Example: A chloride analysis gave the following results:

Sample No.	% Cl^-	Deviation from mean (d)	Deviation from median (dm)
X_1	24.39	0.077	0.03
X_2	24.19	0.123	0.17
X_3	24.36	0.047	0.00

$\bar{X} = \frac{72.84}{3} = 24.313$ $w = 24.39 - 24.19 = 0.20$

$\bar{X}_m = 24.36$

5.2.6 Average deviation from mean ($\bar{d}$):

It is an average of the sum of deviations from mean and it is given a symbol of $\bar{d}$.

$\bar{d}$ from the above example $= \frac{0.077+0.123+0.047}{3} = 0.082$

5.2.7 Average deviation from median ($\overline{dm}$):

It is an average of the sum of deviations from median and it is given the symbol of $\overline{dm}$.

$\bar{d}m$ from the above example = $\frac{0.03+0.17+0.00}{3}$ = 0.067

5.2.8 Standard deviation:

It is the square root of the sum of squares of the individual deviation from the mean divided by the total number of measurements in the set.

The resulting values of standard deviations are positive and symbolised as S when the measurements are few (less than twenty) and δ (sigma) when the measurements are large (more than twenty).

$S = \sqrt{\frac{\sum (xi - \bar{x})^2}{N-1}}$ **When the set of measurements is less than 20.**

$\delta = \sqrt{\frac{(xi - \mu)^2}{N}}$ **When the set of measurements is more than 20.**

$\bar{x}$ is the mean when the set of measurements is less than 20.

μ is the mean when the set of measurements is more than 20.

5.2.9 The variance:

It is the square of standard deviation and symbolised as S^2 and δ^2 .

Example: A volume of liquid is measured by six students and they get the following data:

24.29, 24.30, 24.31, 24.32, 24.38, 24.45 ml. Calculate the mean, median, range, standard deviation and variance for these data.

The solution:

The mean ($\bar{x}$) = $\frac{24.29+24.30+24.31+24.32+24.38+24.45}{6}$ =24.34

The data	$\bar{x}$	d	d^2
24.29		0.05	0.0025
24.30		0.04	0.0016
24.31		0.03	0.0009
24.32	24.34	0.02	0.0004
24.38		0.04	0.0016
24.45		0.11	0.0121
			Total =0.0191

$$S=\sqrt{\frac{\sum(xi-\bar{x})^2}{N-1}} = \sqrt{\frac{0.0191}{6-1}} = 0.0618$$

Therefore, the standard deviation =0.0618

The variance =S^2 =$(0.0618)^2$ =0.0038

The median = $\frac{24.31+24.32}{2} = \frac{48.63}{2} = 24.315$

The range =24.45 - 24.29 = 0.16 .

5.3 The accuracy:

It means the nearness or closeness of the experimental results to the accepted or actual value. As the observed value is more close to the accepted value, the result is more accurate. The term accuracy expresses the extent of the approaching of the result to the real value, while the term precision expresses the extent of the measurements (data of the experiment) from each other.

Methods of expression the accuracy:

The accuracy is always described in terms of errors.

5.3.1 Absolute error (E_a): **It is the difference between the observed value and the accepted value.**

E_a =O-A E_a = absolute error.

O = observed value

A = accepted value

The sign of the absolute error should be mentioned.

5.3.2 Relative error: It is expressed as percentage or in part per thousand (ppt) of the accepted value.

$$\text{Relative error } (E_r)\% = \frac{O-A}{A} \times 100$$

$$\text{or Relative error } (E_r) = \frac{O-A}{A} \times 1000 \text{ as ppt}$$

Example: The accepted value of Cl in NaCl is equal 60.7%. Two students have determined this percentage. The first student has determined the %Cl by volumetric analysis which = 60.1 %. The second student has determined the %Cl by gravimetric analysis which =60.5 %. Which percentage is more accurate for the analysis of Cl in NaCl?

The solution:

$$\text{Relative error} = \frac{60.1 - 60.7}{60.7} \times 100 = -1.00\%$$

$$\text{Relative error} = \frac{60.5 - 60.7}{60.7} \times 100 = -0.33\%$$

The result of second student is more accurate since the relative error is less and his result is more close to the accepted value.

5.4 Significant figures:

They are the confident figures which are considered for calculation. The figures are always approximate and they become more accurate when the used instrument is more precised. For example, if the weight of substance = 15.8 g, the significant figures are three. If the weight is measured by more accurate balance to get 15.81 g or 15.79 g, the weight is expressed as follows

15.80 ± 0.01 and then we get four significant figures.

The significant figures include all the figures involving the uncertain figures. For example: 0.000161, 0.00161, 0.0161, 0.161, 1.61 and 16.1 have three significant figures. For the value 15.315 there are five significant figures.

The significant figures are written in a scheme to indicate the accuracy of measurements. Thus, the weight 4.3062 g is approximated to nearest 0.1mg (0.0001g) which has five significant figures.

5.4.1 The Zeros and the significant figures

Special attention should be paid to whether zeros are significant or not. The following cases are important.

1- **The initial zeros to the right of the decimal point are always significant. For examples:**
   28.0 ml has three significant figures.
   28.00 ml has four significant figures.
   28.000 has five significant figures.
2- **Zeros on the right of decimal and before the digit are not significant and their advantage is to locate the position of the decimal. Examples:0.000851, 0.00851, 0.0851, 0.851, 8.51 and 85.1 are all having three significant figures.**
3- **The zeros that appear between the digit in numbers or logs are always significant. Examples: 1907, 1097, 7908 and 7098 have four significant figures.**
4- **Terminal zeros are significant and insignificant. Ex: If the number of students of Al-Anbar University is 10000, therefore all zeros on the right are significant. But if the number of populations of Ramadi city is million 1000,000, therefore it has only four significant figures, since**

there is uncertainty in the other zeros, because of the difficulty in controlling the exact number of people. Therefore, if the number of population of Ramadi city is written as 1000×10^3, it means that the number is known for approximate thousand of persons. Avogadro's number is 6.023×10^{23} which has four significant figures and the distance between sun and earth is 9300×10^4 miles, which shows also four significant figures.

5.4.2 Rules of Using Significant Figures.

1- **In the calculations and results it is recommended to keep a number of significant figures which include only one uncertain number.**

2- **To apply the rule (1), the values are rounding off to the wanted significant figures by the following steps:**

a- **If the digit on the right of wanted figure is less than five, it is discarded and the figure stays as it is: 3.1413 is rounded off to 3.141 .**

b- **If the digit on the right of wanted figure is more than five, it is rounded off to one and added to the wanted figure: 3.1416 is rounded off to 3.142 .**

c- **If the digit on the right of wanted figure is five, there are here two cases:**

i) **If the wanted figure is even, the approximation is rejected: 1.345 or 1.3450 is approximated 1.34 .**

ii) **If the wanted figure is odd, then it is approximated to one and added to the wanted figure: 1.335 or 1.3350 $\longrightarrow$ 1.34 .**

5.4.3 Addition and Subtraction.

The result after addition and subtraction is approximated to keep the fewest digits on the right of decimal. Examples:

$$\begin{array}{r} 25.34 \\ 5.465 \\ +\ 0.322 \\ \hline 31.127 \end{array} \xrightarrow{\text{approxtd.}} 31.1 \qquad \begin{array}{r} 25.0 \\ 0.0038 \\ +\ 0.00001 \\ \hline 25.00381 \end{array} \xrightarrow{\text{approxtd.}} 25.0 \qquad \begin{array}{r} 4.20 \\ 1.6523 \\ +\ 0.015 \\ \hline 5.8673 \end{array} \xrightarrow{\text{approxtd.}} 5.87$$

5.4.4 Multiplication and Division

The answer is approximated after multiplication and division to a figure which contains the lowest significant figure. Examples:

$7.485 \times 8.61 = 64.4458 \longrightarrow 64.4$.

$0.1642 \div 1.52 = 0.108$, **It is usually approximated.**

$0.0233 \times 22.4 \times 0.092 = 12.4121984 \longrightarrow 12.4$.

5.5 Problems:

1- If you are given the following data: 2.500, 2.504, 2.511, 2.512 and 2.498. Calculate the mean and median of this data and find their difference.

2- A sample of copper was analysed and the following percentages are obtained: 24.8%, 24.93% and 24.69, and the accepted value is 25.059%. Calculate the mean of this copper analysis and find the absolute and relative errors of the analysis. Express your calculations in percentages and in ppt.

3- A limestone contains 30.12% CaO and 2.69% Fe_2O_3 . An analyst has obtained the following percentages: 30.36% CaO and 2.61% Fe_2O_3 . Calculate the absolute and relative errors of this analysis.

4- Nine chemists have analysed samples of protein and they got the following percentages: 35.01%, 34.86%, 34.92%, 35.36%, 35.11%, 35.10%, 34.77%, 35.19& and 34.98%. Calculate the mean, median, the average deviation from mean, the average deviation from median, the standard deviation and variance of these results.

5- A standard sample contains 1.29% water. This sample is analysed by student A who got the following percentages: 1.28, 1.26, and 1.29%. Another student B has analysed another sample contains water in the percentage 8.67, but he got the following percentages of water: 8.48, 8.55 and 8.53%.

   a) Compare the absolute and relative deviations from mean for the two sets of data.

   b) Compare the absolute and relative errors for the two sets of data from mean.

6- Nine Herroin samples has been analysed by Gas chromatography. Every sample is analysed twice to get the following results: Calculate and compare the standard deviations of this analytical method.

8.24	**8.40**	**8.60**	**8.90**	**8.30**	**9.07**	**8.40**	**9.40**	**7.80**
8.27	**8.70**	**8.50**	**8.60**	**8.20**	**9.02**	**8.80**	**9.80**	**7.90**

7- Locate significant figures of the following values:

0.00012, 2.7005, 3.5700, 6.28×10^{-5}g, 10.10g, 31848, 9004, 996, 0.0607, 1.69×10^{-5}, 9.00×10^{3} and 123.45 .

8- Carry out the following additions and subtractions, and approximate your results to the convenient significant figures:

a) 12.67 + 357.8 + 1.349 .

b) 0.0025 + 2.5×10^{-3} + 0.1025 .

c) 0.543 + 543 + 54.3 + 543 .

d) 35.6982 – 13.4397 .

e) 107.97 – 23 .

f) 1.634 + 029 .

g) $6.53\times10^{-2} - 6.53\times10^{-3}$.

9- Carry out the following multiplications and divisions and approximate your results to the convenient significant figures:

(a) $\dfrac{3.50\times0.1503}{35.07\times0.562}$ (b) $\dfrac{5.735\times0.565}{27.40\times6.8164}$ (c) $\dfrac{25.67\times0.1123}{1.023\times0.5000}$

(d) $\dfrac{3.5\times10^{-6}\times7.66\times10^{-6}}{35\times0.253}$ (e) $\dfrac{5.43\times543}{0.543\times54.3}$