

جامعة الانبار

كلية : الصيدلة

قسم : الكيمياء الصيدلانية

اسم المادة باللغة العربية: الكيمياء التحليلية

اسم المادة باللغة الإنكليزية: **Analytical Chemistry**

المرحلة: الاولى

التدريسي: نغم خيري كريم

عنوان المحاضرة باللغة العربية: تقييم دقة البيانات التحليلية

عنوان المحاضرة باللغة الإنكليزية: **The Evaluation of the Reliability of**

Analytical Data

2.1 Introduction

Every physical measurement is subject to a degree of uncertainty that, at best, can be decreased only to an acceptable level. The determination of the magnitude of this uncertainty is often a difficult task that requires additional effort, ingenuity, and good judgment on the part of the scientist.

A direct relationship exists between the accuracy of an analytical measurement and the time and effort spend.

2.2 The Mean and Median

The mean, arithmetic mean, and average (\bar{x}) 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.

The median of a set is that result about which all others are equally distributed, half being numerically greater and half being numerically smaller. If the set consists of an odd number of measurements, selection of the median may be made directly; for a set containing an even number of measurements, the average of the central pair is taken.

Example 1

Calculate the mean and the median for 10.06, 10.20, 10.08, 10.10.

$$\text{mean} = (\bar{x}) = \frac{10.06 + 10.20 + 10.08 + 10.10}{4} = 10.11$$

Since the set contain an even number of measurements, the median is the average of the middle pair:

$$\text{median} = \frac{10.08 + 10.10}{2} = 10.09$$

Ideally, the mean and the median should be numerically identical, however, this condition is not realized, particularly when the number of measurements in the set is small.

2.3 Precision

The term precision is used to describe the reproducibility of results. It can be defined as the agreement between the numerical values of two or more measurements that have been made in an identical fashion. Several methods exist for expressing the precision of data.

* *Absolute methods for expressing precision*

The deviation from the mean ($x_i - \bar{x}$) is a common method for describing precision and is simply the numerical difference, without regard to sign, between an experimental value and the mean for the set of data that include the value. Precision can also be reported in terms of deviation from the median. To illustrate, suppose that a chloride analysis yielded the following results:

Sample	Percent Chloride	Deviation from Mean ($x_i - \bar{x}$)	Deviation from Median
x_1	24.39	0.077	0.03
x_2	24.19	0.123	0.17
x_3	24.36	0.047	0.00
	3[72.94	3[0.247	3[0.20

$$(\bar{x}) = 24.313 = 24.31 \quad \text{Avg} = 0.082 = 0.08 \quad \text{Avg} = 0.067 =$$

0.07

$$w = x_{\max} - x_{\min} = 24.39 - 24.19 = 0.20$$

The spread or range (w) in a set of data is the numerical difference between the highest and lowest result and also a measure of precision. In the previous example, the spread would be 0.20% chloride.

Two other measures of precision are the standard deviation and the variance.

* *Relative precision*

It is often more convenient to indicate the precision relative to the mean (or the median) in terms of percentage or as parts per thousand. For example, for sample x_1 ,

$$\text{relative deviation from mean} = \frac{0.077 \times 100}{24.31} = 0.32 \approx 0.3\%$$

Similarly, the average deviation of the set from the median can be expressed as:

$$\text{relative average deviation from median} = \frac{0.067 \times 1000}{24.36} = 3 \text{ ppt}$$

2.4 Accuracy

The term accuracy denotes the nearness of a measurement to its accepted value and is expressed in terms of error. Accuracy involves a comparison with respect to a true or accepted value; in contrast, precision compares a result with other measurements in the same way.

The accuracy of a measurement is often described in terms of absolute error, which can be defined as:

$$E = x_i - x_t$$

The absolute error E is the difference between the observed value x_i and the accepted value x_r . The value may itself be subjected to considerable uncertainty.

Returning to the previous example, suppose that the accepted value for the percentage of chloride in the sample is 24.36%. The absolute error of the mean is thus $24.31 - 24.36 = -0.05\%$ chloride; here we ordinarily retain the sign of the error to indicate whether the result is high or low.

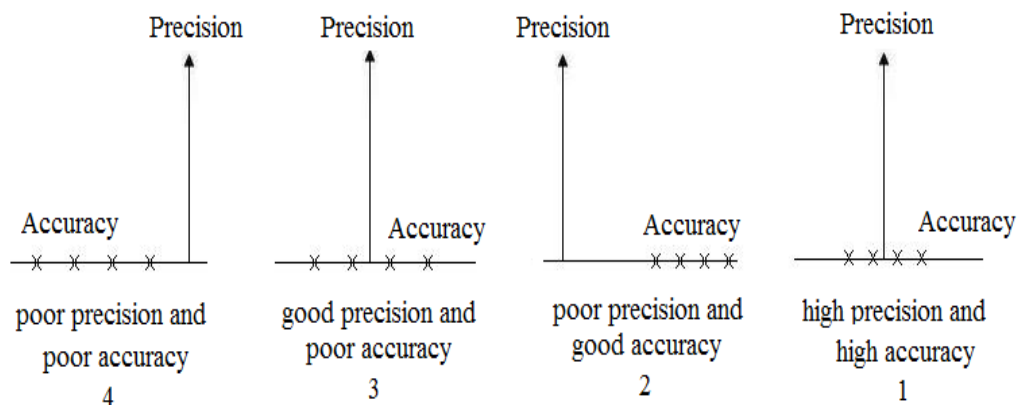
Often a more useful quantity than the absolute error is the relative error expressed as a percentage or in part per thousand of the accepted value. Thus, for the chloride analysis we have been considering,

$$\text{relative error} = \frac{0.05 \times 100}{24.36} = -0.21 = 0.2\%$$

$$\text{relative error} = \frac{0.05 \times 1000}{24.36} = -2.1 = -2 \text{ ppt}$$

2.5 Precision and Accuracy of Experimental Data

The precision of a measurement is readily determined by replicate experiments performed under identical conditions. Unfortunately, an estimate of the accuracy is not equally available because this quantity requires sure knowledge of the vary information that is being sought, namely, the true value. It is tempting to ascribe a direct relationship between precision and accuracy. The danger of this approach is illustrated in the following figure:



2.6 Classification of Errors

The uncertainties that arise in a chemical analysis, and that are responsible for the behavior illustrated in the last figure, may be classified into two broad categories, depending upon their origin.

* **Determinate errors:** those that have a definite value which can (in principle, if not in practice) be measured and accounted for.

* **Indeterminate errors:** result from extending a system of measurement to its maximum. These errors cannot be positively identified and do not have a definite measurable value; instead, they fluctuate in a random manner.

2.7 Types of Determinate Errors

1- Personal errors

These errors are the result of the ignorance, carelessness, prejudices, or physical limitations of the experimenter. For example, they may arise from the use of an improper technique in transferring a sample, from disregard of temperature corrections for a measuring device, from over- or underwashing a precipitate, or from transposing numbers when recording an experimental measurement.

2- Instrumental errors

Instrumental errors are attributable to imperfections in the tools with which the analyst works or to the effects of environmental factors upon these tools. For example, volumetric equipment such as burets, pipets, and volumetric flasks frequently deliver or contain volumes slightly different from those indicated by their graduations, particularly when they are employed at temperatures which differ significantly from the temperature at which they were calibrated. Calibration at the proper temperature will obviously eliminate this type of determinate error.

3- Method errors

Determinate errors are often introduced from nonideal chemical behavior of the reagents and reactions upon which the analysis is based. Such sources of nonideality include the slowness of some reactions, the incompleteness of others, the lack of stability of some species, the nonspecificity of most reagents, and the possible occurrence of side reactions which interfere with the measurement process. For example, in a gravimetric analysis the chemist is confronted with the problem of isolating the element to be determined as a solid of the greatest possible purity. If he fails to wash it sufficiently, the precipitate will be contaminated with foreign substances and have a spuriously high weight. On the other hand, sufficient washing to remove these contaminants may cause weighable quantities to be lost owing to the solubility of the precipitate; here, a negative determinate error will result.

2.8 Effect of Determinate Error upon the Results of an Analysis

Determinate errors generally fall into either of two categories, *constant* or *proportional*. The magnitude of a constant error stays essentially the same as the size of the quantity measured is varied.

Proportional errors increase or decrease according to the size of the sample taken for analysis.

* **Constant errors.** For any given analysis, a constant error will become more serious as the size of the quantity measured decreases. This problem is illustrated by the solubility losses that attend the washing of a precipitate.

Example 2

Suppose that directions call for washing of the precipitate with 200 ml of water, and that 0.50 mg is lost in this volume of wash liquid. If 500 mg of precipitate are involved, the relative error due to solubility loss will be:

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

Loss of the same quantity from 50 mg of precipitate will result in a relative error of – 1.0%.

The amount of reagent required to bring about the color change in a volumetric analysis is another example of constant error. This volume, usually small, remains the same regardless of the total volume of reagent required. Again, the relative error will be more serious as the total volume decreases. Clearly, one way of minimizing the effect of constant error is to use as large a sample as is consistent with the method at hand.

* **Proportional errors.** Interfering contaminants in the sample, if not eliminated in some manner, will cause an error of the proportional variety. For example, a method widely employed for the analysis of copper involves reaction of the copper(II) ion with potassium iodide; the quantity of iodine produced in the reaction is then measured. Iron(III), if present, will also liberate iodine from potassium iodide.

Unless steps are taken to prevent this interference, the analysis will yield erroneously high results for the percentage of copper since the iodine produced will be a measure of the sum of the copper and iron in the sample. The magnitude of this error is fixed by the fraction of iron contamination and will produce the same relative effect regardless of the size of sample taken for analysis. If the sample size is doubled, for example, the amount of iodine liberated by both the copper and the iron contaminant will also be doubled.

2.9 Effects of Indeterminate Error

As suggested by its name, indeterminate error arises from uncertainties in a measurement that are unknown and cannot be controlled by the scientist. The effect of such uncertainties is to produce a scatter of results for replicate measurements.

2.10 The standard deviation

The absolute standard deviation, s , describes the spread of individual measurements about the mean and is given as:

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}$$

where x_i is one of n individual measurements, and \bar{x} is the mean. Frequently, the relative standard deviation, s_r , is reported. The percent relative standard deviation is obtained by multiplying s_r by 100%.

$$s_r = \frac{s}{\bar{x}}$$

Example 3

Calculate the standard deviation s for a subset consisting of the following values of the replicate measurements from the calibration of a 10 ml pipette.

Trial	Vol. water delivered, ml
1	9.990
2	9.986
3	9.973
4	9.983
5	9.980

x_i	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
9.990	7.6×10^{-3}	57.8×10^{-6}
9.986	3.6×10^{-3}	13.0×10^{-6}
9.973	9.4×10^{-3}	88.4×10^{-6}
9.983	0.6×10^{-3}	0.4×10^{-6}
9.980	2.4×10^{-3}	5.8×10^{-6}
5 49.912		
9.9824 = \bar{x}		$\sum_{i=1}^5 (x_i - \bar{x})^2 = 165.4 \times 10^{-6}$

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad s = \sqrt{\frac{165.4 \times 10^{-6}}{5 - 1}} = 6.4 \times 10^{-3} = \pm 0.006$$

We can also calculate s_r :

$$s_r = \frac{s}{\bar{x}} \quad s_r = \frac{0.006}{9.982} = 6.010$$

2.11 The Variance

The variance is another common measure of spread; it is the square of the standard deviation. The standard deviation, rather than the variance, is usually reported because the units for standard deviation are the same as that for the mean value.

To calculate the variance of the last example:

$$\text{Variance} = s^2 = (0.006)^2 = 0.000036$$

2.12 Rejection of Data

When a set of data contains an outlying result that appears to differ excessively from the average (or the median), the decision must be made to retain or to reject it.

Of the numerous statistical criteria suggested to aid in deciding whether to retain or reject a measurement, the Q test is to be preferred. Here the difference between the questionable result and its nearest neighbor is divided by the spread of the entire set. The resulting ratio, Q , is then compared with rejection values that are critical for a particular degree of confidence. Table 2-1 provides critical values of Q at the 90% confidence level. Q = rejection quotient, Q_{crit} = critical Q -value, Q_{exp} = experimental Q -value.

Table 2-1 Critical Values for Rejection Quotient Q

Number of Observations	Q_{crit} (90% confidence) Reject if $Q_{\text{exp}} >$
2	-
3	0.94
4	0.76
5	0.64
6	0.56
7	0.51
8	0.47

9	0.44
10	0.41

Example 4

The analysis of a calcite sample yielded CaO percentage of 55.95, 56.00, 56.04, 56.08, and 56.23, respectively. The last value appears anomalous; should it be retained or rejected?

The difference between 56.23 and 56.08 is 0.15. The spread (56.23 – 55.95) is 0.28. Thus,

$$Q_{exp} = \frac{0.15}{0.28} = 0.54$$

For five measurements, Q_{crit} is 0.64. Since $0.54 < 0.64$, retention is indicated.