

Errors in Chemical Analysis

DR. SUJIT K. BERA

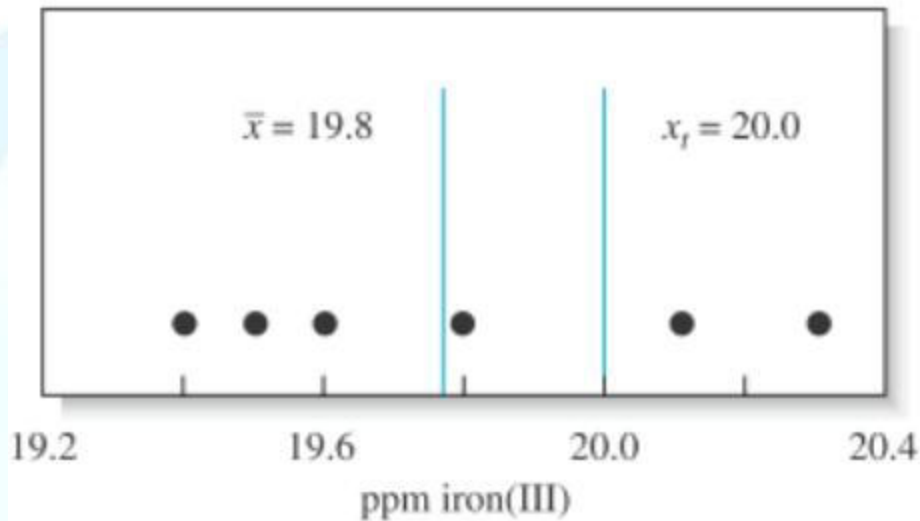
- Every measurements invariably involve some errors and uncertainties.
- it is impossible to perform a chemical analysis that is totally free of errors or uncertainties
- We can only try to minimize errors and estimate their size with acceptable accuracy
- Errors are caused by faulty calibrations or standardizations or by random variations and uncertainties in results.
- Frequent calibrations, standardizations, and analyses of known samples can sometimes be used to lessen all but the random errors and uncertainties.

The term error has two slightly different meanings.

- 1) error refers to the difference between a measured value and the “true” or “known” value.
- 2) error often denotes the estimated uncertainty in a measurement or experiment.

“ We can only hope to minimize errors and estimate their size with acceptable accuracy”

Every measurement is influenced by many uncertainties, which combine to produce a scatter of results.



The figure shows a results from six replicate determinations of iron in aqueous samples of a standard solution containing 20.0 ppm iron(III). Note that the results range from a low of 19.4 ppm to a high of 20.3 ppm of iron. The average, or mean value, \bar{x} , of the data is 19.78 ppm, which rounds to 19.8 ppm

Because measurement uncertainties can never be completely eliminated, *measurement data can only give us an estimate of the “true” value.*

However, the probable magnitude of the error in a measurement can often be evaluated. It is then possible to define limits within which the true value of a measured quantity lies with a given level of probability.

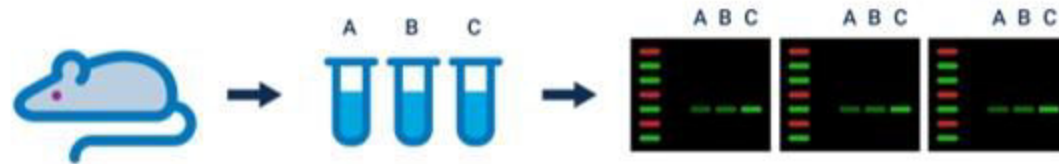
Before beginning an analysis ask, **“What maximum error can be tolerated in the result?”**

The answer to this question often determines the method chosen and the time required to complete the analysis.

Some important terms

To improve the reliability and to obtain information about the variability of results, **two to five portions (replicates)** of a sample are usually carried through an entire analytical procedure.

Replicates are samples of about the same size that are carried through an analysis in exactly the same way.



- Individual results from a set of measurements are seldom the same
- Usually, the **“best” estimate** is considered to be **the central value** for the set.
- The central value of a set should be more reliable than any of the individual results.
- Usually, the **mean or the median** is used as **the central value** for a set of replicate measurements.

An analysis of the variation in the data allows us to estimate the uncertainty associated with the central value.

The Mean and the Median

The **mean**, also called the arithmetic mean or the average, is obtained by dividing the sum of replicate measurements by the number of measurements in the set:

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$$

- The symbol $\sum x_i$ means to add all of the values x_i for the replicates; x_i represents the individual values of x making up the set of N replicate measurements.
- The **median** is the middle value in a set of data that has been arranged in numerical order.
- The median is used advantageously when a set of data contain **an outlier**. An **outlier** is a result that differs significantly from others in the set.
- An **outlier** can have a significant effect on the mean of the set but has no effect on the median.

EXAMPLE 5-1

Calculate the mean and median for the data shown in Figure 5-1.

Solution

$$\text{mean} = \bar{x} = \frac{19.4 + 19.5 + 19.6 + 19.8 + 20.1 + 20.3}{6} = 19.78 \approx 19.8 \text{ ppm Fe}$$

Because the set contains an even number of measurements, the median is the average of the central pair:

$$\text{median} = \frac{19.6 + 19.8}{2} = 19.7 \text{ ppm Fe}$$

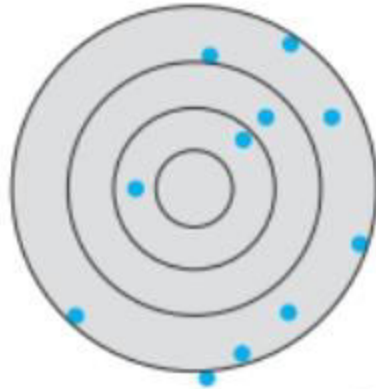
Precision

- * Precision describes the agreement among several results obtained in the same way. Describes the reproducibility of measurements.
- * Precision is readily determined by simply repeating the measurement on replicate samples.
- * Precision of a set of replicate data may be expressed as standard deviation, variance, and coefficient of variation.
- * d_i , deviation from mean, is how much x_i , the individual result, deviates from the mean.

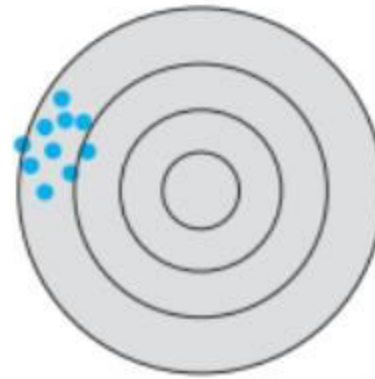
$$d_i = |x_i - \bar{x}|$$

Accuracy

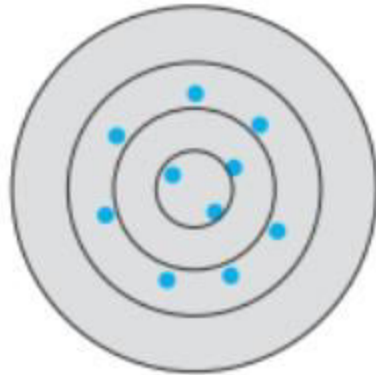
- Accuracy indicates the closeness of the measurement to the true or accepted value and is expressed by the error.
- Accuracy measures agreement between a result and the accepted value.
- Accuracy is often more difficult to determine because the true value is usually unknown. An accepted value must be used instead.
- Accuracy is expressed in terms of either absolute or relative error.



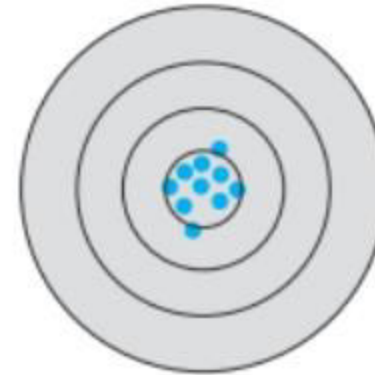
Low accuracy, low precision



Low accuracy, high precision



High accuracy, low precision



High accuracy, high precision

Note that we can have very precise results (upper right) with a mean that is not accurate and an accurate mean (lower left) with data points that are imprecise.

Absolute Error

* The absolute error of a measurement is the difference between the measured value and the true value. If the measurement result is low, the sign is negative; if the measurement result is high, the sign is positive.

$$\text{Absolute Error} = |v_A - v_E|$$

$$\text{Percentage Error} = \left| \frac{v_A - v_E}{v_E} \right| \times 100\%$$

$v_A = \text{approximate (measured) value}$
 $v_E = \text{exact value}$

Types of Errors in Experimental Data

- Results can be precise without being accurate and accurate without being precise.
- Each dot represents the error associated with a single determination.
- Each vertical line labeled $(\bar{x}_i - x_t)$ is the absolute average deviation of the set from the true value.

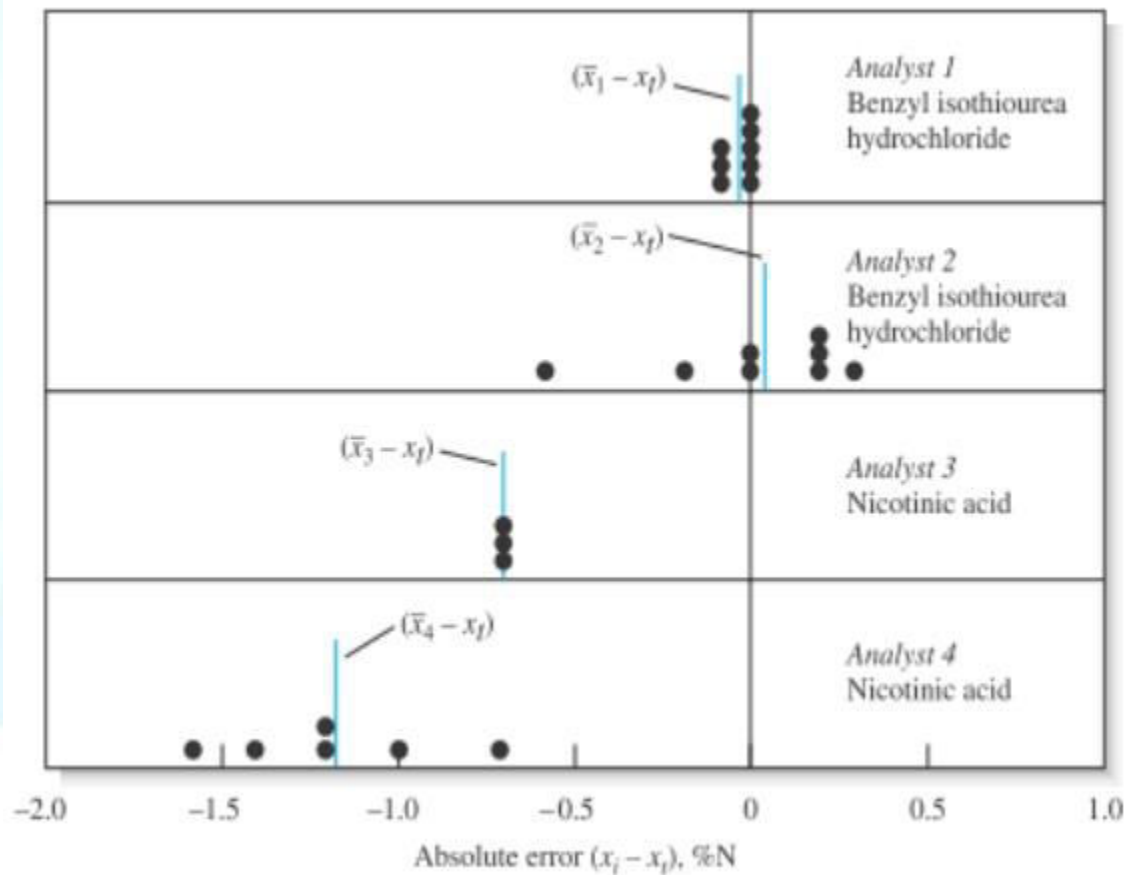


Figure 5-3 Absolute error in the micro-Kjeldahl determination of nitrogen.

Relative Error

- More useful
- Expressed as %, ppt, ppm relative to the sample size

$$\text{Relative Error (E}_r\text{)} = \frac{(X_i - X_t)}{X_t} \times 100; \quad \text{In \%}$$
$$= \frac{(X_i - X_t)}{X_t} \times 1000; \quad \text{In PPT}$$

In general

$$E_r = \frac{x_i - x_t}{x_t} \times 10^n$$

n= 2,3,6.....

- *It is independent upon the size of sample*
- *It is generally expressed in percentage or parts per thousand.*
- *Relative error also possesses the sign.*

Chemical analyses are affected by at least two types of errors:

1. Systematic (or determinate) error, causes the mean of a data set to differ from the accepted value.
2. Random (or indeterminate) error, causes data to be scattered more or less symmetrically around a mean value.

A third type of error is **gross error**.

- These differ from indeterminate and determinate errors.
- They usually occur only occasionally, are often large, and may cause a result to be either high or low.
- They are often the product of human errors.
- Gross errors lead to outliers, results that appear to differ markedly from all other data in a set of replicate measurements.

Systematic errors:

- Systematic errors
 - have a definite value,
 - an assignable cause, and
 - are of the same magnitude for replicate measurements made in the same way.
 - They lead to bias in measurement results.

There are three types of systematic errors:

- Instrumental errors.
- Method errors
- Personal errors

Instrumental Errors

- ✓ are caused by nonideal instrument behavior, by faulty calibrations, or by use under inappropriate conditions
- ✓ Pipets, burets, and volumetric flasks may hold or deliver volumes slightly different from those indicated by their graduations.
- ✓ Calibration eliminates most systematic errors of this type.
- ✓ Electronic instruments can be influenced by noise, temperature, pH and are also subject to systematic errors.
- ✓ Errors of these types usually are detectable and correctable.

Method Errors

- ❑ The nonideal chemical or physical behavior of the reagents and reactions on which an analysis is based often introduce systematic method errors.
- ❑ Such sources of nonideality include the slowness of some reactions, the incompleteness of others, the instability of some species, the lack of specificity of most reagents, and the possible occurrence of side reactions that interfere with the measurement process.
- ❑ Errors inherent in a method are often difficult to detect and hence, these errors are usually the most difficult to identify and correct.

Personal Errors

- result from the carelessness, inattention, or personal limitations of the experimenter.
- Many measurements require personal judgments.
- Examples include estimating the position of a pointer between two scale divisions, the color of a solution at the end point in a titration, or the level of a liquid with respect to a graduation in a pipet or buret.
- Judgments of this type are often subject to systematic, unidirectional errors.
- A universal source of personal error is prejudice, or bias.
 - Number bias is another source of personal error that varies considerably from person to person.
 - The most frequent number bias encountered in estimating the position of a needle on a scale involves a preference for the digits 0 and 5.
 - Also common is a prejudice favoring small digits over large and even numbers over odd.
 - Digital and computer displays on pH meters, laboratory balances, and other electronic instruments eliminate number bias because no judgment is involved in taking a reading.

The Effect of Systematic Errors on Analytical Results

➤ Systematic errors may be either *constant or proportional*.

➤ *Constant Errors*

- ❑ The magnitude of a constant error stays essentially the same as the size of the quantity measured is varied.
- ❑ With constant errors, the absolute error is constant with sample size, but the relative error varies when the sample size is changed

EXAMPLE 5-2

Suppose that 0.50 mg of precipitate is lost as a result of being washed with 200 mL of wash liquid. If the precipitate weighs 500 mg, the relative error due to solubility loss is $-(0.50/500) \times 100\% = -0.1\%$. Loss of the same quantity from 50 mg of precipitate results in a relative error of -1.0% .

- ❑ One way of reducing the effect of constant error is to increase the sample size until the error is acceptable.
- ❑ The excess of reagent needed to bring about a color change during a titration is another example of constant error.
- ❑ This volume, usually small, remains the same regardless of the total volume of reagent required for the titration. Again, the relative error from this source becomes more serious as the total volume decreases.
- ❑ One way of reducing the effect of constant error is to increase the sample size

➤ *Proportional Errors*

- ❑ Proportional errors decrease or increase in proportion to the size of the sample.
- ❑ A common cause of proportional errors is the presence of interfering contaminants in the sample.
- ❑ For example, a widely used method for the determination of copper is based on the reaction of copper(II) ion with potassium iodide to give iodine. The quantity of iodine is then measured and is proportional to the amount of copper. Iron(III), if present, also liberates iodine from potassium iodide. Unless steps are taken to prevent this interference, high results are observed for the percentage of copper because the iodine produced will be a measure of the copper(II) and iron(III) in the sample.
- ❑ The size of this error is fixed by the *fraction* of iron contamination, which is independent of the size of sample taken. If the sample size is doubled, for example, the amount of iodine liberated by both the copper and the iron contaminant is also doubled. Thus, the magnitude of the reported percentage of copper is independent of sample size.

Detection and Elimination of Systematic (Instrumental and Personal) Errors:

1. Periodic calibration of equipment is always desirable because the response of most instruments changes with time as a result of component aging, corrosion, or mistreatment.
2. Most personal errors can be minimized by careful, disciplined laboratory work.
3. It is a good habit to check instrument readings, notebook entries, and calculations systematically.
4. Errors due to limitations of the experimenter can usually be avoided by carefully choosing the analytical method or using an automated procedure.

Detection of Systematic (Method) Errors:

- Bias in an analytical method is particularly difficult to detect.
- The best way to estimate the bias of an analytical method is by analyzing Standard reference materials (SRMs).

Analysis of Standard Samples

- The overall composition of a synthetic standard material must closely approximate the composition of the samples to be analyzed.
- Great care must be taken to ensure that the concentration of analyte is known exactly.
- A synthetic standard may not reveal unexpected interferences so that the accuracy of determinations may not be known.

Independent Analysis

- If standard samples are not available, a second independent and reliable analytical method can be used in parallel.
- The independent method should differ as much as possible from the one under study.
- This practice minimizes the possibility that some common factor in the sample has the same effect on both methods.
- Again, a statistical test must be used to determine whether any difference is a result of random errors in the two methods or due to bias in the method under study.

Blank Determinations

- A blank contains the reagents and solvents used in a determination, but no analyte.
- Often, many of the sample constituents are added to simulate the analyte environment, which is called the sample matrix.
- In a blank determination, all steps of the analysis are performed on the blank material. The results are then applied as a correction to the sample measurements.
- Blank determinations reveal errors due to interfering contaminants from the reagents and vessels employed in the analysis.
- Blanks are also used to correct titration data for the volume of reagent needed to cause an indicator to change color.

Variation in Sample Size

As the size of a measurement increases, the effect of a constant error decreases. Thus, constant errors can often be detected by varying the sample size.

RANDOM / INDETERMINATE ERROR

- **Uncontrolled variables in the measurement**
- **Can be positive or negative**
- **Cannot be corrected for**
- **Random errors are independent of each other**

Random errors can be reduced by:

- **Better experiments (equipment, methodology, training of analyst)**
- **Large number of replicate samples**

Random errors show Gaussian distribution for a large number of replicates

Can be described using statistical parameters