3-3 Types of Error

- Every measurement has some uncertainty, which is called experimental error.
- Conclusions can be expressed with a high or a low degree of confidence, but never with complete certainty.
- Experimental error is classified as either systematic or random.

Systematic Error

- Systematic error, also called determinate error, arises from a flaw in equipment or the design of an experiment.
- If you conduct the experiment again in exactly the same manner,

→ the error is reproducible.

→ In principle, systematic error can be discovered and corrected, although this may not be easy.

- For example,

→ a pH meter that has been standardized incorrectly produces a systematic error.

- Suppose you think that the pH of the buffer used to standardize the meter is 7.00, but it is really 7.08.

→ Then all your pH readings will be lowered by 0.08 pH units

→ When you read a pH of 5.60, the actual pH of the sample is 5.68.

→ This systematic error could be discovered by using a second buffer of known pH to test the meter.

- A key feature of systematic error is that it is reproducible.

→ Systematic error may always be positive in some regions and always negative in others.

→ With care and cleverness, you can detect and correct a systematic error.

Random Error

- Random error, also called indeterminate error, arises from the effects of uncontrolled (and maybe uncontrollable) variables in the measurement.

→ Random error has an equal chance of being positive or negative.

→ It is always present and cannot be corrected.

For examples,

- There is random error associated with reading a scale.

→ One person reading the same instrument several times might report several different readings.

- Another random error results from electrical noise in an instrument.

→ Positive and negative fluctuations occur with approximately equal frequency and cannot be completely eliminated

Precision and Accuracy

- **Precision** describes the **reproducibility** of a result.

- If you measure a quantity several times and the values agree closely with one another,

→ your measurement is precise.

- If the values vary widely,

→ your measurement is not precise.

Precision and Accuracy

- **Accuracy** describes **how close** a measured value is to the “true” value.

- If a known standard is available (such as a Standard Reference Material), accuracy is how close your value is to the known value

- A measurement might be reproducible, but wrong.

→ If you made a mistake preparing a solution for a titration, you might do a series of reproducible titrations but report an incorrect result because the concentration of the titrating solution was not what you intended.

→ In this case, the **precision** is good but the **accuracy** is poor.

Absolute and Relative Uncertainty

- Absolute uncertainty expresses the margin of uncertainty associated with a measurement.
- If the estimated uncertainty in reading a calibrated buret is ± 0.02 ml

→ we say that ±0.02 ml is the absolute uncertainty associated with the reading.

- Relative uncertainty compares the size of the absolute uncertainty with the size of its associated measurement.
- The relative uncertainty of a buret reading of 12.35 ± 0.02 ml is a dimensionless quotient:

Relative uncertainty:

$$\boxed{\text{Relative uncertainty} = \frac{\text{absolute uncertainty}}{\text{magnitude of measurement}}}$$

$$= \frac{0.02\ \text{mL}}{12.35\ \text{mL}} = 0.002$$

- The percent relative uncertainty is simply:

Percent relative uncertainty:

$$\boxed{\text{Percent relative uncertainty} = 100 \times \text{relative uncertainty}}$$
$$= 100 \times 0.002 = 0.2\%$$

- If the absolute uncertainty in reading a buret is constant at ± 0.02 ml,

→ the percent relative uncertainty is 0.2% for a volume of 10 mL and 0.1% for a volume of 20 mL.

3-4 Propagation of Uncertainty from Random Error

- We can usually estimate or measure the random error associated with a measurement, such as the length of an object or the temperature of a solution.

- The uncertainty might be based on how well we can read an instrument or on our experience with a particular method.

- If possible, uncertainty is expressed as the standard deviation or as a confidence interval, which are discussed in Chapter 4.

- This section applies only to random error.

- For most experiments, we need to perform arithmetic operations on several numbers, each of which has a random error.

- The most likely uncertainty in the result is not simply the sum of the individual errors,

→ because some of them are likely to be positive and some negative.

→ We expect some cancellation of errors.

Addition and Subtraction

- Suppose you wish to perform the following arithmetic, in which the experimental uncertainties, designated e_1, e_2, and e_3 are given in parentheses.

$$
\begin{array}{rl}
1.76\ (\pm 0.03) & \leftarrow e_1 \\
+\ 1.89\ (\pm 0.02) & \leftarrow e_2 \\
-\ 0.59\ (\pm 0.02) & \leftarrow e_3 \\
\hline
3.06\ (\pm e_4) &
\end{array}
$$

- The arithmetic answer is 3.06.

→ But what is the uncertainty associated with this result?

- For addition and subtraction, the uncertainty in the answer is obtained from the absolute uncertainties of the individual terms as follows:

Uncertainty in addition and subtraction:

$$e_4 = \sqrt{e_1^2 + e_2^2 + e_3^2}$$

$$
\begin{array}{rl}
1.76\ (\pm 0.03) & \leftarrow e_1 \\
+\ 1.89\ (\pm 0.02) & \leftarrow e_2 \\
-\ 0.59\ (\pm 0.02) & \leftarrow e_3 \\
\hline
3.06\ (\pm e_4) &
\end{array}
$$

Uncertainty in addition and subtraction:

$$\boxed{e_4 = \sqrt{e_1^2 + e_2^2 + e_3^2}}$$

$$e_4 = \sqrt{(0.03)^2 + (0.02)^2 + (0.02)^2} = 0.04_1$$

- The absolute uncertainty e_4 is ± 0.04

→ we express the answer as 3.06 ± 0.04

- Although there is only one significant figure in the uncertainty,

→ we wrote it initially as with the first insignificant figure subscripted.

→ we retain one or more insignificant figures to avoid introducing round-off errors into later calculations through the number

$$\begin{array}{rl} 1.76\ (\pm 0.03) & \leftarrow e_1 \\ +\ 1.89\ (\pm 0.02) & \leftarrow e_2 \\ \underline{-\ 0.59\ (\pm 0.02)} & \leftarrow e_3 \\ 3.06\ (\pm e_4) & \end{array}$$

Uncertainty in addition and subtraction:

$$\boxed{e_4 = \sqrt{e_1^2 + e_2^2 + e_3^2}}$$

$$e_4 = \sqrt{(0.03)^2 + (0.02)^2 + (0.02)^2} = 0.04_1$$

- To find the percent relative uncertainty in the sum, we write

$$\text{Percent relative uncertainty} = \frac{0.04_1}{3.06} \times 100 = 1._3\%$$

- The uncertainty, 0.04_1 is $1._3$ % of the result, 3.06.

→ The subscript 3 in $1._3$% is not significant.

- When we express the final result,

$$\begin{array}{ll} 3.06\ (\pm 0.04) & \text{(absolute uncertainty)} \\ 3.06\ (\pm 1\%) & \text{(relative uncertainty)} \end{array}$$

Multiplication and Division

- For multiplication and division,

→ first convert all uncertainties into percent relative uncertainties.

→ Then calculate the error of the product or quotient as follows:

Uncertainty in multiplication and division:

$$\%e_4 = \sqrt{(\%e_1)^2 + (\%e_2)^2 + (\%e_3)^2}$$

- For example, consider the following operations:

$$\frac{1.76\ (\pm 0.03) \times 1.89\ (\pm 0.02)}{0.59\ (\pm 0.02)} = 5.64 \pm e_4$$

- First convert absolute uncertainties into percent relative uncertainties.

$$\frac{1.76\ (\pm 1._7\%) \times 1.89\ (\pm 1._1\%)}{0.59\ (\pm 3._4\%)} = 5.64 \pm e_4$$

$$\frac{1.76\ (\pm 1._7\%) \times 1.89\ (\pm 1._1\%)}{0.59\ (\pm 3._4\%)} = 5.64 \pm e_4$$

- Then find the percent relative uncertainty of the answer by using Equation 3-6.

$$\%e_4 = \sqrt{(\%e_1)^2 + (\%e_2)^2 + (\%e_3)^2}$$

$$\%e_4 = \sqrt{(1._7)^2 + (1._1)^2 + (3._4)^2} = 4._0\%$$

- The answer is $5.6_4\ (\pm 4._0\%)$

- To convert relative uncertainty into absolute uncertainty,

→ find $4._0$ % of the answer.

$$4._0\% \times 5.6_4 = 0.04_0 \times 5.6_4 = 0.2_3$$

$$4._0\% \times 5.6_4 = 0.04_0 \times 5.6_4 = 0.2_3$$

- The answer is $5.6_4\ (\pm 0.2_3)$.

- Finally, drop the insignificant digits.

$$5.6\ (\pm 0.2) \qquad \text{(absolute uncertainty)}$$
$$5.6\ (\pm 4\%) \qquad \text{(relative uncertainty)}$$

→ The denominator of the original problem, 0.59, limits the answer to two digits.

$$\frac{1.76\ (\pm 0.03) \times 1.89\ (\pm 0.02)}{0.59\ (\pm 0.02)} = 5.64 \pm e_4$$

Mixed Operations

- Now consider a computation containing subtraction and division:

$$\frac{[1.76\ (\pm 0.03)\ -\ 0.59\ (\pm 0.02)]}{1.89\ (\pm 0.02)} = 0.619_0 \pm ?$$

1) Work out the difference in the numerator, using absolute uncertainties.

$$\sqrt{(0.03)^2 + (0.02)^2} = 0.03_6.$$

$$1.76\ (\pm 0.03) - 0.59\ (\pm 0.02) = 1.17\ (\pm 0.03_6)$$

2) Convert into percent relative uncertainties

$$\frac{1.17\ (\pm 0.03_6)}{1.89\ (\pm 0.02)} = \frac{1.17\ (\pm 3._1\%)}{1.89\ (\pm 1._1\%)} = 0.619_0\ (\pm 3._3\%)$$

$$\sqrt{(3._1\%)^2 + (1._1\%)^2} = 3._3\%.$$

- The percent relative uncertainty is $3._3$%

  → so the absolute uncertainty is $0.03_3 \times 0.619_0 = 0.02_0$

- The final answer can be written as

$$0.619_0\ (\pm 0.02_0) \qquad \text{(absolute uncertainty)} \qquad \Rightarrow \qquad 0.619\ (\pm 0.02)$$

$$0.619_0\ (\pm 3._3\%) \qquad \text{(relative uncertainty)} \qquad \Rightarrow \qquad 0.619\ (\pm 3\%)$$

- *(Caution)* The result of a calculation ought to be written in a manner consistent with its uncertainty.

→ Because the uncertainty begins in the 0.01 decimal place,

it is reasonable to round the result to the 0.01 decimal place:

$$0.62\ (\pm 0.02) \qquad \text{(absolute uncertainty)}$$

$$0.62\ (\pm 3\%) \qquad \text{(relative uncertainty)}$$

The Real Rule for Significant Figures

- The real rule: The first uncertain figure is the last significant figure.
  → The first digit of the absolute uncertainty is the last significant digit in the answer.

- For example, in the quotient

$$\frac{0.002\,364\,(\pm 0.000\,003)}{0.025\,00(\pm 0.000\,05)} = \cancel{0.09456\ (\pm\ 0.000\ 2)}$$

$$= 0.094\,6\,(\pm 0.000\,2)$$

→ the uncertainty (± 0.000 2) occurs in the fourth decimal place.

→ even though the original data have four figures,
the answer 0.094 6 is properly expressed with three significant figures

$$\frac{0.821\ (\pm 0.002)}{0.803\ (\pm 0.002)} = 1.022\ (\pm 0.004)$$

- Even though the dividend and divisor each have three figures,

  → The quotient is expressed with four figures

- The quotient 82/80 is better written as 1.02 than 1.0, if we do not know its uncertainty.

  → The actual uncertainty lies in the second decimal place, not the first decimal place, if uncertainties are in ones place

  → If I write 1.0,

  you can surmise that the uncertainty is at least $1.0 \pm 0.1 = \pm 10\%$

- Therefore, when an answer lies between 1 and 2,

  → It is all right to keep one extra digit

Exponents and Logarithms

- For the function $y = x^a$,

  → the percent relative uncertainty in y ($\%e_y$) is equal to ***a*** times the percent relative uncertainty in x ($\%e_x$)

Uncertainty for powers and roots:

$$y = x^a \Rightarrow \%e_y = a(\%e_x)$$

- For the function $y = x^{1/2}$,

  → a 2% uncertainty in x will yield a (1/2)(2%) = 1% uncertainty in y.

- If $y = x^2$,

  → a 3% uncertainty in x leads to a (2)(3%) = 6% uncertainty in y

- If y is the base 10 logarithm of x,

  → then the absolute uncertainty in y (e_y) is proportional to the relative uncertainty in x (e_x/x):

Uncertainty for logarithm:

$$y = \log x \quad \Rightarrow \quad e_y = \frac{1}{\ln 10}\frac{e_x}{x} \approx 0.434\,29\frac{e_x}{x}$$

- Now consider y = antilog x, which is the same as saying $y = 10^x$

  → the relative uncertainty in y is proportional to the absolute uncertainty in x.

Uncertainty for 10^x:

$$y = 10^x \quad \Rightarrow \quad \frac{e_y}{y} = (\ln 10)e_x \approx 2.302\,6\, e_x$$

- Table 3-1 summarizes rules for propagation of uncertainty.
  → You need not memorize the rules for exponents, logs, and antilogs, but you should be able to use them.

See Table 3-1

Analytical Chemistry

Chapter 4. Statistics

- Standard deviation and error curve
- Confidence interval
- Student *t*
- The method of least squares
- Calibration curves

Statistics

- All measurements contain experimental error,

  → so it is never possible to be completely certain of a result.

  → Statistics gives us tools to accept conclusions that have a high probability of being correct and to reject conclusions that do not

Count on “normal” days	Today’s count
5.1 5.3 4.8×10^6 cells/μL 5.4 5.2	5.6×10^6 cells/μL

- “Is my red blood cell count today higher than usual?”
- If today’s count is twice as high as usual,

  → it is probably truly higher than normal.

- But what if the “high” count is not excessively above “normal” counts?

  → To scientifically answer the question, we need statistics

4-1. Gaussian Distribution

- If an experiment is repeated a great many times and if the errors are purely random,

  → the results tend to cluster symmetrically about the average value

- The more times the experiment is repeated,

  → the more closely the results approach an ideal smooth curve

  → the Gaussian distribution.

See Fig 4-1

Mean Value and Standard Deviation

- In the hypothetical case,
  → A manufacturer tested the lifetimes of 4768 electric light bulbs.
  → The bar graph shows the number of bulbs with a lifetime in each 20-h interval.
- Because variations in the construction of light bulbs, such as filament thickness and quality of attachments, are random,
  → Lifetimes approximate a Gaussian distribution that best fits the data.

See Fig 4-1

- Light bulb lifetimes, and the corresponding Gaussian curve, are characterized by two parameters.

1) **Arithmetic mean** (also called the average)

→ the sum of the measured values (x_i) divided by n, the number of measurements:

Mean: $$\bar{x} = \frac{\sum_i x_i}{n}$$

2) **Standard deviation**, s,

→ measures how closely the data are clustered about the mean.

→ a measure of the uncertainty of individual measurements

→ the smaller the standard deviation,

the more closely the data are clustered about the mean

Standard deviation: $$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}}$$

(n-1): the degrees of freedom

Standard Deviation and Probability

- The formula for a Gaussian curve is

$$y = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-(x-\mu)^2/2\cdot\sigma^2}$$

x : values of individual measurements

μ : mean for an infinite set of data

(the population mean)

x-μ : deviation from the mean

y : frequency of occurrence for each value of x-μ

σ : standard deviation for an infinite set of data

(the population standard deviation)

See Fig 4-3

$\frac{1}{\sigma \cdot \sqrt{2\pi}}$: normalization factor

→ which guarantees that the area under the entire curve is unity

$$y = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-(x-\mu)^2 / 2 \cdot \sigma^2}$$

- It is useful to express deviations from the mean value in multiples, z, of the standard deviation

See Fig 4-3

$$z = \frac{x - \mu}{\sigma} \approx \frac{x - \bar{x}}{s}$$

$$y = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-z^2 / 2}$$

- The probability of measuring z in a certain range is equal to the area of that range.
- For example, the probability of observing z between -2 and -1 is 0.136.

The standard deviation measures the width of the Gaussian curve

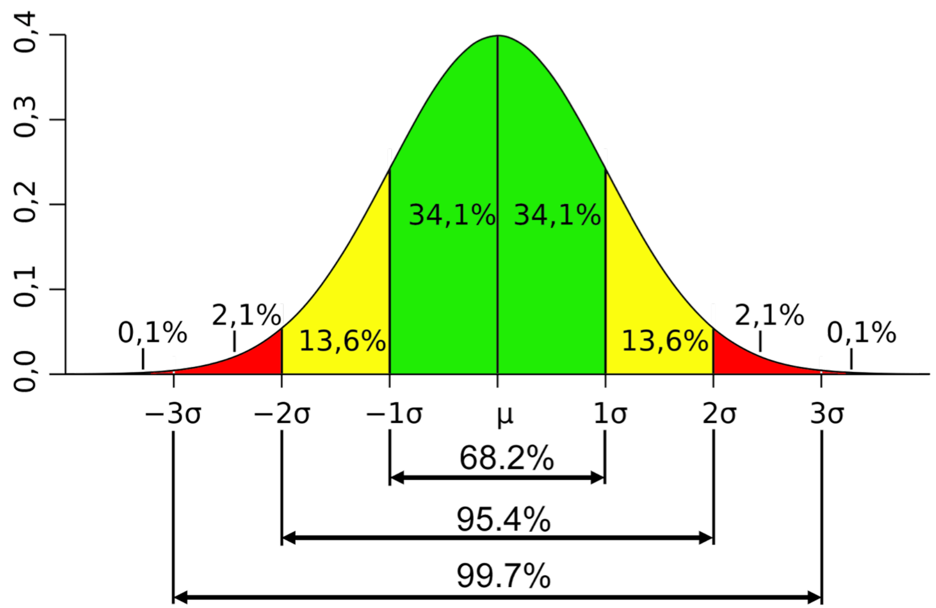
- In any Gaussian curve,

  → 68.3% of the area is in the range from μ - 1σ to $\mu + 1\sigma$

  → more than two-thirds of the measurements are expected to lie within one standard deviation of the mean.

- 95.5% of the area lies within $\mu \pm 2\sigma$
- 99.7% of the area lies within $\mu \pm 3\sigma$

Range	Percentage of measurements
$\mu \pm 1\sigma$	68.3
$\mu \pm 2\sigma$	95.5
$\mu \pm 3\sigma$	99.7



- The mean gives the center of the distribution.

- The standard deviation measures the width of the distribution

→ The larger the value of s, the broader the curve.

See Fig 4-2

- An experiment that produces a small standard deviation is more precise than one that produces a large standard deviation.

- Greater precision does not necessarily imply greater accuracy, which means nearness to the "truth."

- Suppose that you use two different techniques to measure sulfur in coal: Method A has a standard deviation of 0.4%, and method B has a standard deviation of 1.1%.

- You can expect that approximately two-thirds of measurements from method A will lie within 0.4% of the mean.

- For method B, two-thirds will lie within 1.1% of the mean.

- You can say that method A is more precise.

- Example: Mean and Standard Deviation

→ Find the average and the standard deviation for 821, 783, 834, and 855.

$$\bar{x} = \frac{821 + 783 + 834 + 855}{4} = 823._2$$

$$s = \sqrt{\frac{(821 - 823.2)^2 + (783 - 823.2)^2 + (834 - 823.2)^2 + (855 - 823.2)^2}{(4 - 1)}} = 30._3$$

- We commonly express experimental results in the form:

→ mean ± standard deviation $= \bar{x} \pm s.$

→ We will retain one or more insignificant digits to avoid introducing round-off errors into subsequent work