Data Collection, Processing, and Analysis

Reference: Philip R. Bevington, D. Keith Robinson: *Data Reduction and Error Analysis* for the Physical Sciences, McGraw Hill, 2003, ISBN 0-07-247227-8

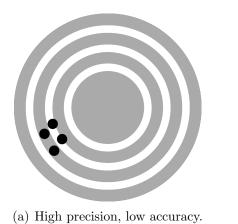
2.1 Error, Precision, Accuracy

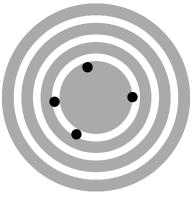
It is natural for one to choose to make a single, careful measurement of a given parameter. By selectively making a single measurement of each parameter in a given equation, one can plug the numbers into an appropriate formula and calculate some end-parameter in question. This approach may work well in most general applications; however, such procedure is generally *unacceptable* in the laboratory. In science, high precision and a good *estimate of the error*, or uncertainty in the determined value, is required. Therefore, multiple measurements of <u>each parameter</u> must be made, multiple calculations of the end value are performed, and the uncertainty in the end-parameter is determined by statistical methods.

2.1.1 Error

Error is the difference between an observed or calculated value and the true value. Usually, the "true" value is unknown (or else we would not be doing the experiment in the first place). We may know approximately what it should be from earlier experiments or theoretical predictions. Errors fall into three classes:

- **Illegitimate errors:** These result from simple mistakes in reading an instrument, or blunders in a calculation. These errors can be corrected by carefully repeating the operations.
- **Random errors** are often beyond the control of the experimenter. Random errors cannot be identified or estimated by a single measurement. The reason for such errors are vast and often depend on the experimental setup and measuring devices. IF the errors are truly random, the data points should be symmetrically distributed on either side





(b) Low precision, high accuracy

Figure 2.1: Precision versus accuracy.

of the actual value, as shown in Figure 2.2 on page 20. Using a sufficient number of data points one can get a good estimation of the *probable value* using the statistical methods outlined in this chapter.

Systematic errors consistently contribute to an overestimation or underestimation of the values. The presence and magnitude of systematic errors can be difficult to detect if the expected value is not known. Such errors may result from faulty calibration, poor experimental technique, or other such reason. The experimenter must be very careful and methodical to avoid the introduction of systematic errors.

2.1.2 Precision

Precision is a measure of the reproducibility of your measurements. Webster's defines precision as *exactness*, but this definition needs to be refined for use in scientifice experiments. We understand precision as a measure of how well the result has been determined, and how reproducable it is.

2.1.3 Accuracy

The laboratory manual defines Accuracy as a measure of how closely the result agrees to the "actual" value. This, of course, requires knowledge of what this "actual" value might be. Webster's considers Accuracy and Precision to be synonymous, and this is a common point of confusion. Accuracy should be thought of more in terms of target shooting. If a person hits the bullseye, we don't naturally call the shot "precise", we call it "accurate" or "on target". Either of these are equally legitimate, grammatically, but it is more natural to refer to "good aim" with the term "accurate". Such is the case here. It is entirely a matter of semantics, but language must be understood to be useful. Results that are "close to the mark" are considered "accurate;" therefore, **Accuracy is determined by the difference between the resultant value and the "accepted" value.**. See Fig. 2.1 for a graphical comparison between "precision" and "accuracy".

A quantitative measure can be made on the Accuracy of a given result. There are several methods which are implemented, depending on the existence or absence of an "accepted value" (what your lab manual calls *actual value*). These are given on page 22 by Equations 2.4.7 and 2.4.8, which are **Percent Error** and **Percent Difference**, respectively. Note that the former, Percent Error, requires a theoretical or generally accepted value with which to compare. Percent Difference is used to compare my results with yours & does not require an "accepted value."

2.2 Measurements and Least Measures

All laboratory equipment share two common factors: Units and Finite Resolution. Good laboratory procedure mandates that **all measurements are recorded to the precision allowed by the instrument**. Several points may be taken on reports where measurements have been arbitrarily reduced in precision by sloppy work. The units of the measurements should be properly recorded in the column where the data is collected. Using tables, and collecting data in columns, reduces the amount of writing and allows for quick visual evaluations of your data. Careful study of columns of data may reveal to your eye some pattern in the behavior of the numbers. One column in a data collection table should be reserved for each measurement; therefore, it is common that all numbers in a column possess the same number of decimal places.

The Least Measure of the equipment, also known as minimum error for reasons which will be covered shortly, is determined by the decimal places indicated on the equipment. Your common meter stick, for example, can be read within half a millimeter; therefore, the *least measure* of the meter stick is 0.05 cm and <u>all readings</u> with a meter stick should be recorded to 2 decimal places when using the centimeter scale. The triple beam balance can be interpolated to 1/100 gram, so the *least measure* of the triple beam balance is 0.01 g and <u>all readings</u> with a triple beam balance should be recorded to 2 decimal places when using the recorded to 2 decimal places when using the network of the triple beam balance is 0.01 g and all readings with a triple beam balance should be recorded to 2 decimal places when using units of grams.

2.3 Working with Significant Figures

A very important issue that arrises in the laboratory is the "significance" of the numbers. Instruments always possess a finite degree of precision, and that precision is reflected in the number of *significant figures* that are given for any measurement. All subsequent calculations must accurately reflect the precision of the various numbers used to obtain that result, thus procedures for handling significant figures are laid forth.

Suppose we want to calculate the value v = d/t where d is the distance measured with a meter stick and t is the time measured with a digital stopwatch. The least measure of the meter stick is 0.05 cm, and the least measure of the stopwatch is 0.001 s. Suppose we measure the following:

$$d = 1.25 \,\mathrm{cm}$$

 $t = 1.378 \,\mathrm{s}$

We have the distance to 3 significant figures and the time to 4 significant figures. What this means is the following: the distance traveled is $1.25 \text{ cm} \pm 0.05 \text{ cm}$ over a time of $1.378 \text{ s} \pm 0.001 \text{ s}$. In other words, we don't know for sure that the distance is 1.25 cm, 1.30 cm, or 1.20 cm. For this reason, the 5 in the distance measurement is called the *least significant figure* and is <u>always</u> assumed to be in error. Similarly the 8 in the time measurement is the least significant figure. It, too, is assumed to be in error. The least significant figure is only a *best estimate* of what this number should be, and that is the reason we cannot go further with this measurement. If the 5 is wrong in the distance, is would be pointless to say what is beyond that position.

When we take the velocity calculation v = d/t we get

$$v = 0.907111756 \,\frac{\mathrm{cm}}{\mathrm{s}}$$

from our calculator. Our meter stick and stop watch told us how accurate the base measurements were, but we have no such guide here. Which decimal place is the limit of our precision?

The product (or quotient) of two or more numbers has the same number of significant figures as the number in the product (or quotient) which possesses the fewest number of significant figures.

In our case, the distance had the fewest number of significant figures at 3, thus the velocity must be rounded to 3 significant figures:

$$v = 0.907 \,\frac{\mathrm{cm}}{\mathrm{s}}$$

Sums and differences work quite differently. Before taking any sum or difference, the numbers must all be written out to the same number of <u>decimal places</u>. Suppose we wanted to take the following sum:

12.1352
1.123
112.10
2.0153

Although the first number has the most significant figures, the third number actually dictates policy since it has the fewest <u>decimal places</u>. Before taking the sum, we must round all numbers up to 2 decimal places.

12.14	
1.12	
112.10	
2.02	
127.38	

Note the sum is 127.38 not the expected 127.37 that would be obtained by summing the original data. This is not uncommon, and the two methods do in fact agree. Remember that the **least significant figure is always assumed to be in error**. In other words, neither of them are completely right, but they agree *within error*.

Counting Significant Figures

In order to properly track significant figures through calculations, you must first be able to determine which numbers are significant in your measurements and which are not. The following rules apply to all measurements:

- All non-zero numbers in a measurement ARE significant.
- Zeroes to the left of non-zero numbers are NOT significant.
- Zeroes to the right of non-zero numbers ARE significant.

For example:

measurement	# sig-figs
12.14	4
01.12	3
0.10	2
2.02	3
127.380	6

That leaves only one final point. Suppose the following number is recorded:

500

This number could have 1, 2, or 3 significant figures. There really is no way of knowing; therefore, this method of writing such numbers is considered **ambiguous** and **bad form**. If a number such as this MUST be written, *scientific notation* is preferred:

 5.0×10^2

which makes it absolutely clear that there are 2, not 1 or 3, significant figures in this measurement.

2.4 Error Analysis

2.4.1 Gaussian Distribution

When measurements of a single parameter are made repeatedly, the result will very likely have a distribution as shown in Figure 2.2. This bell-shaped curve is called the normal, or Gaussian, distribution. It describes the probability of obtaining a particular measured value. If one assumes that the error is truly random, then for a large set of measurements the frequency is likely distributed evenly on either side of the actual value; therefore, the actual value will likely be near the center of the distribution, or the *mean value* of the distribution. It follows that the mean value is the *best estimate* of the measured parameter and will be treated as the *determined value* of the parameter.

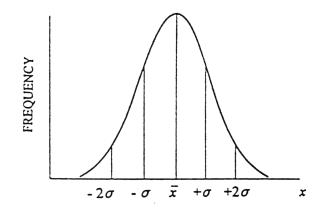


Figure 2.2: A typical distribution arising from random errors.

2.4.2 Mean and Standard Deviation

When multiple measurements are made of a single parameter, the set of data collected is referred to as the sample data, or simply the sample. The number of points in the sample is referred to as the sample size N. Consider a set of 6 measurements of a length yielding the results $x_1 = 2.15$, $x_2 = 2.12$, $x_3 = 1.96$, $x_4 = 2.22$, $x_5 = 2.06$, and $x_6 = 2.09$, all in millimeters. These might represent measurements with a micrometer gauge whose minimum length resolution is 0.01 mm. Since the random errors are larger than the length resolution, we adopt the following definition: The "best value" of the measurements is the **arithmetic mean** and is calculated as

$$\bar{x} = \frac{2.15 + 2.12 + 1.96 + 2.22 + 2.06 + 2.09}{6}$$
 mm = 2.10 mm.

(Sometimes, the notation $\langle x \rangle$ is used instead of \bar{x} .) This can be represented in a symbolic way as

$$\bar{x} = \frac{1}{N} (x_1 + x_2 + x_3 + x_4 + x_5 + x_6)$$

= $\frac{1}{N} \sum_{i=1}^{N} x_i$, (2.4.1)

where the symbol \sum represents a sum. Next, we define the deviation d_i of each value from the mean by

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

where \bar{x} is the mean value from Eq. (2.4.1). Some of these deviations are positive, and some are negative. It can easily be shown that the average deviation, defined by

$$\bar{d} = \frac{1}{N} \sum_{i=1}^{N} d_i , \qquad (2.4.2)$$

is zero. Of more use is the **root-mean-square (rms) deviation** σ . To estimate σ , we square each deviation d_i (remember that squares are always ≥ 0), add them together, divide by N - 1, and take the square root:

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N d_i^2} \,. \tag{2.4.3}$$

In the above example, the individual deviations are $d_1 = 0.05$, $d_2 = 0.02$, $d_3 = -0.14$, $d_4 = 0.12$, $d_5 = -0.04$, and $d_6 = -0.01$, all in millimeters. Hence,

$$\sigma_x = \sqrt{\frac{1}{6-1} \left[0.05^2 + 0.02^2 + (-0.14)^2 + 0.12^2 + (-0.04)^2 + (-0.01)^2 \right]} \text{ mm}$$

= 0.0879 mm.

 σ_x tells us how much "scatter" we have in our experiment, i.e. by how much a typical measurement deviates from the mean \bar{x} . From the formula one can easily see that the standard deviation has the same units as the data. Not only that, but the standard deviation of a column of data should be represented to the same number of significant figures as the data in the column.

How good an estimate is our calculated mean value \bar{x} to the (unknown) **true** value x_{true} ? To find out, we'd have to do the above experiment (take 6 time measurements and calculate \bar{x}) over and over again. Each experiment would yield a slightly different \bar{x} , but we expect them to scatter around x_{true} . The statistical quantity used to describe the precision of a *set* of measurements is the **standard deviation of the mean**:

$$\sigma_{x,m} = \frac{\sigma_x}{\sqrt{N}}$$
(2.4.4)
= $\sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N d_i^2}$.

We can decrease $\sigma_{x,m}$ (and hence increase our precision) simply by increasing the number of measurements N. However, as N enters into Eq. (2.4.5) with its square root, we need four times the amount of data to cut $\sigma_{x,m}$ in half. Often, this sets a practical limit as to how precise we can measure. In our example with N = 6 and $\sigma_x = 0.0879$ mm, we obtain $\sigma_{x,m} = 0.0359$ mm.

A more rigorous treatment would further show us that the probability of the true value lying in the range

$\bar{x} \pm \sigma_{x,m}$	is	68.3%
$\bar{x} \pm 2\sigma_{x,m}$	is	95.4%
$\bar{x} \pm 3\sigma_{x,m}$	is	99.7%

Frequently, the standard error or uncertainty δx is taken as twice the standard deviation of the mean:

$$\delta x = 2\sigma_{x,m} = \frac{2\sigma_x}{\sqrt{N}} \tag{2.4.5}$$

The result of our measurement should then be reported in the *standard form* as

$$x = \bar{x} \pm \delta x \,. \tag{2.4.6}$$

In the above example, this would be

$$x = 2.10 \,\mathrm{mm} \pm 0.07 \,\mathrm{mm}$$

or

$$x = (2.10 \pm 0.07) \,\mathrm{mm}$$

It is ok to use 1σ - or 3σ -uncertainties (i. e., $\delta x = \sigma_{x,m}$ or $\delta x = 3\sigma_{x,m}$), as long as this is clearly stated in the lab report. In this lab, however, we prefer to use the 2σ uncertainties.

2.4.3 Significant Figures and Instrument Resolution

As we have seen, the expression of σ , and thus δx , yields only an approximate value based on a finite sampling of experimental values. It follows, therefore, that **only one or at most two**¹ significant figures should be retained in expressions for σ and δx . Since δx indicates the range of validity for the mean \bar{x} , we will round off \bar{x} to the same number of decimal places as contained in the rounded δx . Some examples will help to illustrate this rule:

$$\begin{array}{l} \bar{x} &= 47.265\\ \delta x &= 4.302 \end{array} \right\} \qquad x = 47 \pm 4 \\ \\ \bar{x} &= 0.07153\\ \delta x &= 0.00048 \end{array} \right\} \qquad x = 0.0715 \pm 0.0005 \\ \\ \bar{x} &= 2.1738\\ \delta x &= 0.0129 \end{array} \right\} \qquad x = 2.174 \pm 0.013 \\ \\ \\ \bar{x} &= 5726.4\\ \delta x &= 121.2 \end{array} \right\} \qquad x = (5.73 \pm 0.12) \times 10^3$$

2.4.4 Comparing your Results

There are three circumstances that you can face as an experimenter: (1) you have an accepted value for your determined quantity, (2) you have some other determined value to compare against, or (3) you have nothing but your own data to consider. Results can be considered or compared for each of these three cases as follows.

Percent Error

The easiest (but extremely rare) situation is when you have an **accepted value** (there are very few "known" values) with which you can compare. This estimation of your accuracy is called *percent error* and is calculated using

Percent Error = $\frac{|\text{determined value} - \text{accepted value}|}{\text{accepted value}} \times 100\%$ (2.4.7)

¹Rule: if the first non-zero number of δx is 1, round to 2 significant figures; otherwise round to one figure.

Percent Difference

When you do not have an accepted value, but you do have another determined value to compare against (perhaps from another research group) you can estimate your accuracy using a *percent difference* calculation.

Percent Difference =
$$\frac{|\text{first value} - \text{second value}|}{(\text{first value} + \text{second value})/2} \times 100\%$$
 (2.4.8)

Internal Comparison

The worst case is when you have neither an accepted value nor another determined value to compare against. In this case you can do an estimate of your accuracy using the *uncertainty*, δx , and your mean, \bar{x} . This is similar to the Percent Error calculation of Eq. (2.4.7).

Relative Error
$$\epsilon_x = \frac{\delta x}{\bar{x}}$$
 (2.4.9)

Relative Percent Error
$$= \frac{\delta x}{\bar{x}} \times 100\% = \epsilon_x \times 100\%$$
 (2.4.10)

We will use the relative error in the next section on error propagation.

2.4.5 Error Propagation

Often an experiment requires the measurement of two or more quantities which are then used to calculate the value of something else. If we know the uncertainties in the measured quantities, what is the uncertainty in the calculated quantity. There are two approaches:

- 1. The **"Tabular Method"**, which is straightforward, but often requires a significant amount of calculations.
- 2. The "Calculus Method", which requires far fewer calculations, but involves some language of calculus, which may be unfamiliar to you. Don't let this bother you. You will become familiar with this language as the semester progresses. In the meantime, the examples given in this manual should get you started.

Tabular Method

In the "Tabular Method", data is collected in the columns of tables. Calculations are done row by row using a complete set of unique measurements each time, and the error estimates and analysis are done on the completed columns, sometimes hundreds or thousands of entries in length. Perhaps 500 measurements are taken of each of 20 physical quantities. From this, analysis is done to calculate 500 results, and error analysis is then carried out on those 500 results. This allows the errors to "propagate" normally through the calculations, and we evaluate that error based on the statistical distribution at the end.

Example: the volume of a cylinder is given by

$$V = \frac{\pi d^2 h}{4}$$

d (cm)	$h (\rm cm)$	$V (\mathrm{cm}^3)$
2.31	3.77	63.2
2.32	3.98	67.3
2.18	3.92	58.5
2.31	3.77	63.2
2.32	3.80	64.2
2.18	3.92	58.5
2.31	3.77	63.2
2.32	3.98	67.3
2.18	3.92	58.5
2.28	3.76	61.4

Table 2.1: Sample data collection and analysis for a right cylinder's volume.

Table 2.2: Mean, standard deviation and 2σ uncertainty of the Cylinder Measurements and Volume Calculations.

	$d(\mathrm{cm})$	h (cm)	$V (\rm cm^3)$
Average	2.27	3.86	62.5
St. Dev.	0.0638	0.0928	3.32
2σ Uncertainty	0.128	0.186	6.64

If we were to measure a cylinder, such as a graduated cylinder or pipet, we might measure the diameter, d, ten times. We would similarly measure the length, h, ten times (the same number must be made for each measured quantity, as you will see in a moment). We would then calculate one volume, V, for each pair d and h. That gives ten volumes, as shown in Table 2.1. We would calculate the Mean and Standard Deviation for both d, h, and Vdirectly from the table of values, as shown in Table 2.2. Our final result would then be written in the standard form shown in Equation 2.4.6:

> $d = 2.27 \,\mathrm{cm} \pm 0.06 \,\mathrm{cm}$ $h = 3.86 \,\mathrm{cm} \pm 0.09 \,\mathrm{cm}$ $V = 62 \,\mathrm{cm}^3 \pm 3 \,\mathrm{cm}^3$

Calculus Method

There are times when the Tabular Method is inconvenient due to the large number of calculations. Fortunately, a little calculus shows us a much faster method of calculating the uncertainties of derived quantities. Suppose we have measured quantities A, B, C, etc., and Table 2.3: Error propagation for some common mathematical operations. Some of the results use the relative error, defined as $\epsilon_Q = \delta Q/\bar{Q}$ in Eq. (2.4.9).

- 1. Addition and/or Subtraction: $Q = A \pm B \pm C$ $\delta Q = \sqrt{(\delta A)^2 + (\delta B)^2 + (\delta C)^2}$
- 2. Multiplication: Q = ABC

$$\epsilon_Q = \sqrt{(\epsilon_A)^2 + (\epsilon_B)^2 + (\epsilon_C)^2}$$
 or $\frac{\delta Q}{\bar{Q}} = \sqrt{\left(\frac{\delta A}{\bar{A}}\right)^2 + \left(\frac{\delta B}{\bar{B}}\right)^2 + \left(\frac{\delta C}{\bar{C}}\right)^2}$

3. Division: Q = A/B

$$\epsilon_Q = \sqrt{(\epsilon_A)^2 + (\epsilon_B)^2}$$
 or $\frac{\delta Q}{\bar{Q}} = \sqrt{\left(\frac{\delta A}{\bar{A}}\right)^2 + \left(\frac{\delta B}{\bar{B}}\right)^2}$

4. Raising to a Power: $Q = kA^n$, where k and n are numbers with no uncertainty.

$$\epsilon_Q = n\epsilon_A$$
 or $\frac{\delta Q}{\bar{Q}} = n\frac{\delta A}{\bar{A}}$

5. Exponential Function: $Q = ke^A$, where k is a number with no uncertainty, and e = 2.718... is the Euler's number, the base of natural logarithms.

$$\frac{\delta Q}{\bar{Q}} = \delta A$$

6. Natural Logarithm: $Q = \ln A$ $\delta Q = \frac{\delta A}{\overline{A}}$

associated uncertainties δA , δB , δC , etc. Furthermore, we wish to calculate the uncertainty in a quantity Q which is a function of A, B, C, etc., namely Q = Q(A, B, C, ...). This uncertainty in Q is given by

$$\delta Q = \sqrt{\left(\frac{\partial Q}{\partial A} \,\delta A\right)^2 + \left(\frac{\partial Q}{\partial B} \,\delta B\right)^2 + \left(\frac{\partial Q}{\partial C} \,\delta C\right)^2 + \dots} \tag{2.4.11}$$

The symbol $(\partial Q/\partial A)$ is called the **partial derivative of** Q with respect to A, and is a measure of how Q changes as A is changed with everything else in the expression held constant. Table 2.3 lists the results for some common functions Q. Given below are some examples for which you should work out the arithmetic yourself. Notice that the results are given in a form which is properly rounded off.

• Q = A + B, where $A = 78 \pm 4$ and $B = 14 \pm 3$.

Then $\delta Q = \sqrt{4^2 + 3^2} = 5$, and $Q = 92 \pm 5$.

• Q = B - A, where $B = 116 \pm 12$ and $A = 94 \pm 5$. Then $\delta Q = \sqrt{12^2 + 5^2} = 13$, and $Q = 22 \pm 13$.

•
$$Q = AB$$
, where $A = 30 \pm 2$ and $B = 4.00 \pm 0.10$.
Then $\epsilon_Q = \frac{\delta Q}{\bar{Q}} = \sqrt{\left(\frac{2}{30}\right)^2 + \left(\frac{0.1}{4.0}\right)^2} = 0.071$, and $Q = 120 \pm 9$.

•
$$Q = B/C$$
, where $B = 748 \pm 6$ and $C = 2.73 \pm 0.04$.
Then $\epsilon_Q = \frac{\delta Q}{\bar{Q}} = \sqrt{\left(\frac{6}{748}\right)^2 + \left(\frac{0.04}{2.73}\right)^2} = 0.017$, and $Q = 274 \pm 5$.

- $Q = 3A^2$, where $A = 50.4 \pm 0.5$. Then $\frac{\delta Q}{\bar{Q}} = 2\left(\frac{0.5}{50.4}\right) = 0.0198$, and $Q = (7.62 \pm 0.15) \times 10^3$.
- Q = A(B C), where $A = 5.3 \pm 0.2$, $B = 10.2 \pm 0.8$ and $C = 2.1 \pm 0.5$. Here, we need to apply propagation rules for addition (rule #1 in Tab. 2.3) and multiplication (rule #2). We first calculate the error for the difference D = B - C = 8.1: $\delta D = \sqrt{(\delta B^2) + (\delta C)^2} = \sqrt{0.8^2 + 0.5^2} = 0.94$. Next, we calculate $\frac{\delta Q}{\bar{Q}} = \sqrt{\left(\frac{0.2}{5.3}\right)^2 + \left(\frac{0.94}{8.1}\right)^2} = 0.122$. Then, $Q = 43 \pm 5$.
- $Q = Ae^B$, where $A = 1.0 \pm 0.3$ and $B = 3.0 \pm 0.2$. Here, we need to apply error propagation rules for multiplication (rule #2) **and** exponentiation (rule # 5).

Then
$$\frac{\delta Q}{\bar{Q}} = \sqrt{\left(\frac{0.3}{1.0}\right)^2 + 0.2^2} = 0.36$$
, and $Q = 20 \pm 7$.