

Introduction to Statistics, Error and Measurement

Throughout the semester we will be making measurements. When you do an experiment, it is important to be able to evaluate *how well you can trust your measurements*. For example, the known value of g , the acceleration due to gravity, is $\approx 9.81 \text{ m/s}^2$, (" \approx " means approximately equal to). If you make a measurement that says $g = 10.1 \text{ m/s}^2$, is that measurement "wrong"? How do you compare that measurement to the known value of g ? Suppose you measure some quantity that is *not* known? You may make a number of measurements, and get several different results. For example, suppose you measure the mass of an object three times, and get three different values, 5 kg, 4.8 kg, and 5.4 kg. Can you evaluate what the *real* mass of the object is from those measurements?

The mathematical tools we will learn in this lab will answer some of these questions. They are some of the most basic methods of statistical analysis; they will allow us to give information about our measurements in a standard, concise way, and to evaluate how "correct" our measurements are. The methods we will cover are used in all areas of science which involve taking any measurements, from popularity polls of politicians, to evaluating the results of a clinical trial, to making precise measurements of basic physical quantities.

Let's start with the basics of the different *kinds* of errors, and how to measure them.

Types of Errors

There are two types of errors encountered in experimental physics: **systematic** errors and **random** errors.

Systematic errors can be introduced

- by the design of the experiment
- by problems with the instruments you are using to take your data
- by your own biases

Consider a very simple experiment designed to measure the dimensions of a particular piece of material precisely. A systematic error could be introduced if the measuring instrument is calibrated improperly. For example, a scale might be set a little too low, so that what *reads* as "zero" is *really* "-1 kg". Everything you measure on the scale will come out one kilogram lighter than it really is. If a particular observer always tends to overestimate the size of a measurement, that would also be a systematic error, but one related to the personal characteristics of the experimenter.

Random errors are produced by unpredictable and uncontrollable variations in the experiment. These can be due to the limits of the precision of the measuring device, or due to the experimenter's inability to make the same measurement in **precisely** the same way each time. Even if systematic errors can be eliminated by good experimental design, there will always be some uncertainty due to random errors. Numerical values measured in experiments are therefore never **absolutely** precise; there is always some uncertainty.

Accuracy and Precision

The **accuracy** of a measurement describes how close the experimental result comes to the actual value. That is, it is a measure of the “correctness” of the result. For example, if two independent experiments give the values 2.717 and 2.659 for e (the base of the natural log), the first value is said to be more accurate because the *actual* value of e is 2.718.....

The **precision** of an experiment is a measure of the *reproducibility* of the result. Suppose you measure the same thing three times. The precision would be a measure of how similar all the measurements are to each other. It is a measure of the magnitude of uncertainty in the result. Suppose one person weighs a cat, and comes up with three different masses each time: 10 kg, 12 kg, and 11 kg. Suppose another person weighs the same cat, and comes up with these three masses: 11.1 kg, 11.5 kg, and 11.3 kg. The second person’s measurement would be said to be more precise. (Both people are likely to be scratched, though.)

Significant Figures

When reading the value of an experimental measurement from a calibrated scale, only a certain number of figures or digits can be obtained or read. That is, only a certain number of figures are significant.

The **significant figures** (sometimes called "**significant digits**") of an experimentally measured value include all the numbers that can be read directly from the instrument scale plus one doubtful or estimated number.

For example, if a ruler is graduated in millimeters (mm), you can use that ruler to estimate a length *up to one tenth of a millimeter*. For example, suppose you make a sequence of measurements of the length of some object using this ruler, and get an average value of 318.9811123 mm. The measurement is only accurate to one decimal place, so you would report the number as 319.0 ± 0.1 mm. The “zero” is shown after the decimal point because that digit is the last significant one. All the other digits are meaningless and do not convey any real information about your measurements.

Data Analysis

Percent Error and Percent Difference

How do you measure the size of an error? The object of some experiments is to measure the value of a well-known quantity, such as g . (You’ll be making this measurement yourselves in next week’s experiment!) The most accurate value of these quantities (measured by teams of dedicated professional scientists!!) is the value given in your textbooks and tables. In making a comparison between the results of your experiment and the accepted value measured with much more precision in specialized laboratories, you want to cite the **percent error**, a measurement of how much your measurement differs from the “official” value. The **absolute difference** between the experimental value E and the accepted value A is written $|E-A|$, where the “|” signs mean absolute value. The **fractional error** is the ratio of this **absolute difference** over the accepted value:

$$\text{Fractional error} = \frac{|E - A|}{A}$$

Usually, people convert this fractional error into a percent, and give the **percent error**:

$$\text{Percent error} = \frac{|E - A|}{A} \times 100\%.$$

Sometimes, you need to compare two equally reliable results when an accepted value is *unknown*. This comparison is the **percent difference**, which is the ratio of the **absolute difference** between experimental results E_1 and E_2 to the **average of the two values**, expressed as a percent:

$$\text{Percent difference} = \frac{|E_1 - E_2|}{\left[\frac{(E_1 + E_2)}{2}\right]} \times 100\%.$$

Mean Value

Even if systematic errors can be eliminated from an experiment, the **mean value of a set of measurements of a quantity, x , is a better estimate of the true value of x than is any single measurement**. For this reason, experiments are often repeated a number of times. If we denote $\langle x \rangle$ as the mean value, and there are N measurements x_i (where i varies from 1 to N), then $\langle x \rangle$ is defined by the following equation:

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

In fact, random errors are distributed according to a Gaussian distribution, which looks like the familiar bell curve. In case you aren't familiar with it, the “ Σ ” symbol we just used is a capital Greek letter “sigma “. In math, it is called a summation sign. It means “add up everything to the right”. So in the equation above, the sigma is a shorthand way of writing

$$\langle x \rangle = \frac{1}{N} (x_1 + x_2 + x_3 + \dots + x_N).$$

Mean Deviation

To obtain the **mean deviation** of a set of N measurements, the **absolute deviations** of $|\Delta x_i|$ are determined; that is

$$|\Delta x_i| = |x_i - \langle x \rangle|$$

The **mean deviation** $\langle \Delta x \rangle$ is then

$$\langle \Delta x \rangle = \frac{1}{N} \sum_{i=1}^N |\Delta x_i|$$

Example 1.1 What are the mean value and mean deviation of the set of numbers 5.42, 6.18, 5.70, 6.01, and 6.32?

Solution:

The mean is

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i = \frac{5.42 + 6.18 + 5.70 + 6.01 + 6.32}{5} = 5.926 = 5.93$$

Note that since the original measurements were only valid to two decimal places, the average can only be valid to two decimal places as well, so we round from 5.926 to 5.93.

The absolute deviations for each measurement are:

$$|\Delta x_1| = |5.42 - 5.93| = 0.51$$

$$|\Delta x_2| = |6.18 - 5.93| = 0.25$$

$$|\Delta x_3| = |5.70 - 5.93| = 0.23$$

$$|\Delta x_4| = |6.01 - 5.93| = 0.08$$

$$|\Delta x_5| = |6.32 - 5.93| = 0.39$$

Then the mean deviation is:

$$\langle \Delta x \rangle = \frac{1}{N} \sum_{i=1}^N |\Delta x_i| = \frac{0.51 + 0.25 + 0.23 + 0.08 + 0.39}{5} = 0.292 = 0.29$$

Usually, people report an experimental measurement as the mean “plus or minus” the mean deviation:

$$E = \langle x \rangle \pm \langle \Delta x \rangle$$

In our example, the value would be reported like this: 5.93 ± 0.29 .

Standard Deviation

Statistical theory states that the **precision** of a measurement can be determined using a quantity called the **standard deviation**, σ (called “sigma”, this is the Greek lower-case “s”). The standard deviation of a distribution of measurements is defined as follows:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \langle x \rangle)^2}$$

The **standard deviation** is a measure of **spread**. If the standard deviation is *small*, then the **spread in the measured values about the mean** is *small*, and so the **precision** in the measurements is *high*. The standard deviation is always positive and has the same units as the measured values.

It can be shown, for a Gaussian distribution, that 69% of the data points will fall within one standard deviation of $\langle \mathbf{x} \rangle$, $(\langle \mathbf{x} \rangle - \sigma) < x_i < (\langle \mathbf{x} \rangle + \sigma)$; 95% are within two standard deviations, and only 0.3% are farther than 3σ from $\langle \mathbf{x} \rangle$. So, for example, if an experimental data point lies 3σ from a theoretical prediction, there is a strong chance that either the prediction is not correct or there are systematic errors which affect the experiment.

Example 1.2 What is the standard deviation of the set of numbers given in Example 1.1?

Solution:

First, find the square of the deviation of each of the numbers

$$\Delta x_1^2 = (5.42 - 5.93)^2 = 0.26$$

$$\Delta x_2^2 = (6.18 - 5.93)^2 = 0.06$$

$$\Delta x_3^2 = (5.70 - 5.93)^2 = 0.05$$

$$\Delta x_4^2 = (6.01 - 5.93)^2 = 0.01$$

$$\Delta x_5^2 = (6.32 - 5.93)^2 = 0.15$$

then

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \langle x \rangle)^2} = \sqrt{\frac{0.26 + 0.06 + 0.05 + 0.01 + 0.15}{5}} = 0.33$$

The result of our measurement of E can also be reported as:

$$E = \langle \mathbf{x} \rangle \pm \sigma$$

Note that all these statements are valid if **only** random errors are present.

Presenting Data

Data Tables

You should always organize the experimental data which you take into data tables, in order to make the data very clear and easy to understand. The data table will generally have two or more columns which present the values of the controlled (independent) and measured (dependent) variables. The table should contain labels for the various columns, **including units and errors**. If some variables are fixed during a particular experiment, the values can be given in the title for the table or in the text; similarly, if the errors are the same for all measurements they can be included in the text or caption. Data tables are usually organized by ascending values of the independent variable.

Example 1.3 For a fixed volume, pressure is proportional to temperature for an ideal gas. A fixed amount (10 moles) of gas is placed in a container having a fixed volume of 1m^3 and the temperature is increased from 0°C to 60°C in increments of 10° . The pressure in Newton/m^2 (Pascal) is measured by an apparatus for which the pressure can be read accurately in 100 pascal units with the 10 Pascal digit estimated. The pressure was measured several times and a statistical error determined for each measurement. The independent variable (temperature) is given in the left column of the table on the next page. The measured quantity (a.k.a. the dependent variable) is given, along with the error, in the right column.

GAS PRESSURE AS A FUNCTION OF TEMPERATURE

$V = 1\text{m}^3$; $n = 10$ moles

Temperature ($^{\circ}\text{C}$)	Pressure $P \pm \Delta p$ (Pascal)
0	22700 ± 20
10	23540 ± 20
20	24330 ± 40
30	25210 ± 30
40	26010 ± 30
50	26820 ± 30
60	27730 ± 50

Graphs

It is often useful to visually display your data in the form of a graph. When you are plotting data in a graph, always make it really clear what you are plotting on each axis! Label the axes and show units. We will show you how to do this in the Excel tutorial in the appendix.

