Experimental Uncertainty

Abstract

This is intended as a brief summary of the basic elements of uncertainty analysis, and a handy reference for laboratory use. It provides some elementary "rules-of-thumb" which are satisfactory for use in an introductory physics laboratory.

References

For thorough derivations and justifications of the material presented in this summary, the student should consult one of the following:

Taylor, Introduction to Error Analysis, Second Edition, University Science Books (1997)
Young, Statistical Treatment of Experimental Data, McGraw-Hill (1960)
Beers, Introduction to the Theory of Error Analysis, Addison-Wesley (1958)

Measurement Uncertainties and Uncertainty Propagation

In most experiments, certain quantities are measured and then other quantities are determined from the measured data. For example, we might measure the length and width of a room, and then determine its area by multiplication. Associated with each measured quantity is a "measurement uncertainty". When a final result is determined from measured quantities, these measurement uncertainties lead to an uncertainty in the derived result. The procedure by which an experimenter determines the uncertainty in a final result from the measurement uncertainties is called "error analysis", or "propagation of uncertainty".

Types of Uncertainty (Absolute, Fractional, Percent, Relative)

The **absolute uncertainty** in a quantity is the actual amount by which the quantity is uncertain, e.g.if **L** = 6.0 ± 0.1 cm, the absolute uncertainty in **L** is 0.1 cm. *Note that the absolute uncertainty of a quantity has the same units as the quantity itself.*

The **fractional uncertainty** is the absolute uncertainty divided by the quantity itself, e.g.if **L** = 6.0 ± 0.1 cm, the fractional uncertainty in **L** is 0.1/6.0 = 1/60. *Note that the units cancel in this division, so that fractional uncertainty is a pure number.*

Percent uncertainty is fractional uncertainty expressed as a percent, i.e. fractional uncertainty multiplied by 100. If **L** = 6.0 ± 0.1 cm, its percent uncertainty is 1.7%.

The word "uncertainty", by itself, normally means "absolute" uncertainty. Fractional or percent uncertainties are both called "relative" uncertainties because they relate the size of the uncertainty to the size of the result itself. While not normally used in the presentation of data or experimental results, relative uncertainties are often useful in intermediate stages of "error analysis" calculations.

***Rule #1**:* *Express the uncertainty in each of your measurements or derived results as an absolute uncertainty, using the same units for the quantity and its associated uncertainty.*

Measurement Uncertainties

The first step in determining the reliability of an experiment is to evaluate the measurement uncertainties associated with each measured quantity. While in many cases, measurement uncertainty is due to the particular instruments used, it may also depend on human limitations of the experimenter or on some randomness in the effect being measured.

As an example of instrumental measurement uncertainty, consider a measurement of the length of this page using a meter-stick. When the zero of the meter stick is placed at the top of the page, the bottom appears to lie between 27.9 and 28.0 cm, although somewhat closer to 27.9 cm. Since the marks on the meter-stick are 0.1 or 0.2 centimeters wide, and placing the zero at the top edge of the paper might be "off" by 0.1 centimeter, the experimenter might feel confident saying that the page's length to the nearest tenth of a centimeter is 27.9 cm, but would certainly feel that specifying the length to the nearest hundredth centimeter is risky. Hence, it would be prudent to state the length as 27.9 ± 0.1 centimeters. This situation illustrates two basic rules-of-thumb:

***Rule #2**:* *Make uncertainty estimates large enough to give yourself a margin of safety. You should feel confident that the "real value" of the measured quantity lies somewhere within the uncertainty range you specify.*

***Rule #3**:* *The measurement uncertainty due to an instrument itself may be assumed equal to the smallest scale division of the instrument.*

A common situation in which measurement uncertainty relates more to human limitation than to instrumental error is the measurement of time with an electronic stopwatch. The watch can measure to .01 second. When a careful experimenter makes five separate measurements of the time for a ball to roll a fixed distance down an inclined plane, the results are likely to look like this: 4.63, 4.69, 4.64, 4.56, and 4.70 seconds. While each time is measured to .01 second, the **range** of values is 4.70 minus 4.56, or 0.14 seconds. In another common experiment, a spring gun fires a projectile horizontally from the edge of a lab table. The horizontal distance from the table's edge to the spot where the projectile hits the floor is easily measured to within ± 1 cm, but due to an inherent randomness in the operation of the spring gun, typical results from a series of five trials using the same apparatus might be: 256, 239, 263, 252, and 253 cm. In both these situations, the best an experimenter can do is take an average of the measurements, namely 4.644 seconds and 252.6 cm, respectively. But from the time range of 0.14 seconds and the distance range of 263 - 239 = 24 cm, the experimenter might reasonably estimate the uncertainty in time to be $0.14/2 = 0.07 \rightarrow 0.1$ seconds and distances uncertainty to be $24/2 = 12 \rightarrow 10$ cm. The value for each would then be written as 4.6 ± 0.1 and 250 ± 10 cm. *Notice that the uncertainties are related in only an approximate way to the ranges of values obtained in a series of 5 measurements, **and** that the uncertainties in both cases make it unreasonable to state the results to more than two "significant figures" (i.e 4.6 and 2.5×10^2).*

***Rule #4**:* *If a measurement is not reproducible within instrumental uncertainty, the measurement uncertainty should be estimated from the range of values obtained in a series of five (or more) trials.*

***Rule #5**:* *No quantity should be stated to more significant figures than are justified by its associated uncertainty.*

Uncertainty Propagation: Worst Case Method

The most straightforward way to find the uncertainty in the final result of an experiment is *worst case error analysis*, a method in which uncertainties are estimated from the difference between the largest and smallest possible values that can be calculated from the data. As an illustration of the method, suppose we measure the length and width of this page to be 27.9 ± 0.1 cm and 21.6 ± 0.1 cm, respectively. Then the most probable value for its area is (27.9)(21.6) = 602.64 sq cm. But the largest possible value ("worst case") consistent with the data is (28.0)(21.7) = 607.6 sq cm. Similarly, the smallest possible area is (27.8)(21.5) = 597.7 sq cm. Since the actual area might be anywhere between these extremes, it is reasonable to state the result as 603 ± 5 sq cm.

In a second example, the mass and volume of a sample of oil are measured as 7.54 ± 0.01 g and 10.0 ± 0.1 cm^3, respectively. The density of this oil is most probably (7.54 g)/(10.0 cm^3) = .754 g/cm^3, but might be as large as (7.55 g)/(9.9 cm^3) = .763 g/cm^3 or as small as (7.53 g)/(10.1 cm^3) = .746 g/cm^3. *Note that in this case we must use the* ***smallest*** *possible volume when calculating the* ***largest*** *possible density.* This range of possible density values suggests that the final result should be expressed as 0.75 ± 0.01 g/cm^3.

Rule #6: *When a result is calculated from several measured quantities, its uncertainty is approximately half the difference between the largest and smallest possible results that could be obtained from the measurements and their associated measurement uncertainties.*

Differential Error Analysis

A second, and somewhat more formal, method of treating the propagation of uncertainties is related to differential calculus. When an experimental result, **f**, depends on several independent measured quantities (**x**, **y**, **z**, ...) in such a way that **f** can be written as a differentiable function of the measured quantities, the change in **f** which results from small changes in the measured values can be determined from the calculus of partial derivatives:

$$f \;=\; f(x, y, z, ...)$$

$$df \;= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz + ...$$

Since an uncertainty can be interpreted as a range within which a quantity is free to change, we can relate the absolute uncertainty in **f** (call it **Δf**) to the absolute uncertainties **Δx**, **Δy**, **Δz**, etc, in a simple way:

Rule #7: *If a final result, **f**, is a differentiable function of several measured quantities (**x**, **y**, **z**, ...), the uncertainty in **f** relates to measurement uncertainties according to:*

$$\Delta f = \left|\frac{\partial f}{\partial x}\right|\Delta x + \left|\frac{\partial f}{\partial y}\right|\Delta y + \left|\frac{\partial f}{\partial z}\right|\Delta z + ...$$

The absolute value signs (along with the fact that x, y, z, etc are positive by definition) guarantee that the contributions of different measurement uncertainties to the uncertainty of the final result are always additive. This is in contrast to the calculus of analytic functions, in which variables can change in such a way as to produce canceling effects on the function itself.

Special Cases--Results of Differential Error Analysis

We now proceed to look at some special cases of Rule #7, in an attempt to find some simple rules-of-thumb:

1) $f = x + y$ $\qquad \dfrac{\partial f}{\partial x} = 1 \qquad \dfrac{\partial f}{\partial y} = 1 \qquad \Delta f = \Delta x + \Delta y$

2) $f = x - y$ $\qquad \dfrac{\partial f}{\partial x} = 1 \qquad \dfrac{\partial f}{\partial y} = -1 \qquad \Delta f = \Delta x + \Delta y$

3) $f = xy$ $\qquad \dfrac{\partial f}{\partial x} = y \qquad \dfrac{\partial f}{\partial y} = x \qquad \Delta f = y\Delta x + x\Delta y$

4) $f = \dfrac{x}{y}$ $\qquad \dfrac{\partial f}{\partial x} = \dfrac{1}{y} \qquad \dfrac{\partial f}{\partial y} = -\dfrac{x}{y^2} \qquad \Delta f = \dfrac{1}{y}\Delta x + \dfrac{x}{y^2}\Delta y$

The $\mathbf{\Delta f}$ in each case is the absolute uncertainty in **f**. For #3, the fractional uncertainty turns out to have a particularly simple form:

$$\frac{\Delta f}{f} = \frac{y}{xy}\Delta x + \frac{x}{xy}\Delta y$$

$$= \frac{\Delta x}{x} + \frac{\Delta y}{y}$$

The fourth example also yields a simple result if we look at the relative error:

$$\frac{\Delta f}{f} = \frac{y}{yx}\Delta x + \frac{xy}{xy^2}\Delta y$$

$$= \frac{\Delta x}{x} + \frac{\Delta y}{y}$$

These four examples yield two simple rules:

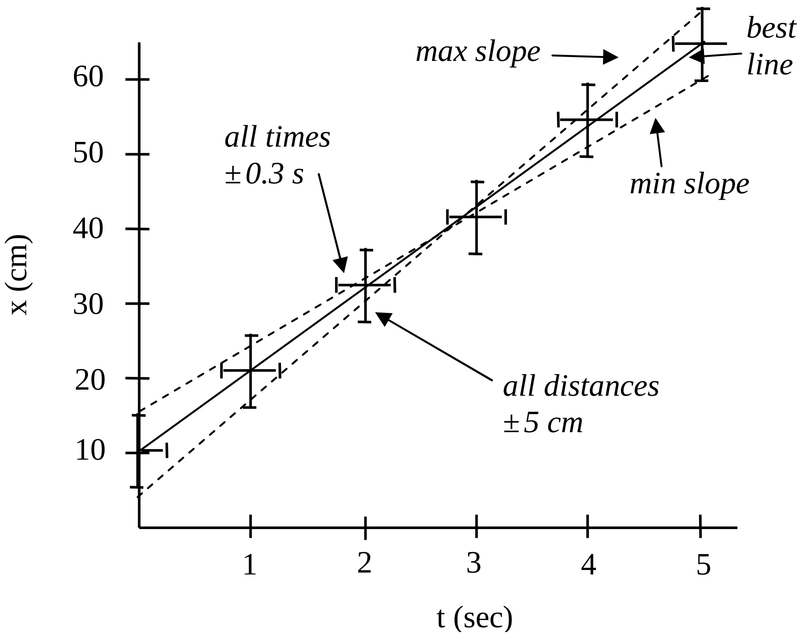
***Rule #8**:* *The absolute uncertainty of the sum or difference of two quantities is the sum of their absolute uncertainties.*

***Rule #9**:* *The relative uncertainty of the product or quotient of two quantities is the sum of their relative uncertainties.*

Graphs and Error Flags - Uncertainty in the Slope of a Line

In a typical experiment to determine velocity, the position of an object is recorded every second. Suppose the uncertainty in each position measurement is ± 5 cm, the uncertainty in each time measurement is ± .25 s, and the data for a five-second interval is given by the table to the left of the following figure.

t(sec)	**x(cm)**
0	10
1	21
2	33
3	42
4	55
5	65



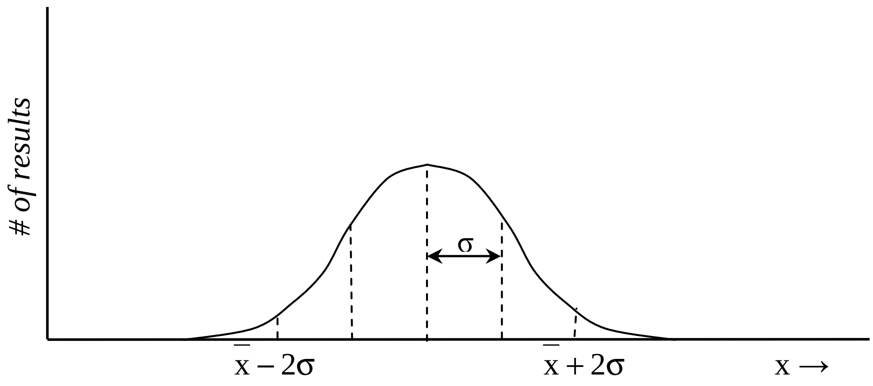
In the graph to the right, the uncertainties in each data point are indicated by "error flags", the vertical flags are for uncertainty in distance, horizontal for time. The lengths of this flags correspond to the uncertainty ranges. For example, since the point at **t** = (3.0 ± 0.3) seconds corresponds to **x** = (42 ± 5) cm, the error flag extends from 37 cm to 47 cm in the vertical direction, and from 2.7 sec to 3.3 sec in the horizontal direction.

The solid line drawn through the data points was estimated by an experimenter to be the "best fit" to this data. The slope of this line, **V** = dx/dt = 11 cm/sec, represents the most probable value of the velocity. But there are many other straight lines which could have been drawn through the error flags, and hence there is a range of possible slopes consistent with the data. To estimate this uncertainty in the slope, we draw the "worst case" lines, i.e.the line of greatest slope and the line of least slope which still pass through most of the error flags. If you think about the slope being given by rise/run, then to get a "max" slope, you will need the "max" rise/"min" run. The worst case lines are shown as dotted lines in the figure, and have slopes of $\mathbf{V}_{\mathbf{max}}$ = 13 cm/sec and $\mathbf{V}_{\mathbf{min}}$ = 9 cm/sec. Given these extremes, it is reasonable to express the final result for the velocity as **V** = (11 ± 2) cm/sec.

***Rule #10**:* *The uncertainty in the slope of a straight line fit to experimental data is approximately half the range of slopes of the set of all straight lines which can be drawn through the data error flags.*

Standard Deviation and Confidence Limits

We have seen that when a measurement is not reproducible because of random errors, its uncertainty can be estimated from the results of several trials. A complete theory of random error predicts that for a **very large** number of trials, **N**, the distribution of results will look like the bell-shaped curve shown in the following figure.



The peak in the curve corresponds to the average result, **x**, and the curve's width is a measure of the range of results obtained, i.e. the uncertainty in the measurement. If **x(i)** is the result of the i^{th} trial, we can define the "deviation" of the i^{th} result as **[x(i)-x]**, and the "standard deviation", **σ** , as:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left[x(i) - \bar{x}\right]^2}$$

The standard deviation is defined in such a way that 68.3% of all results lie between **x - σ** and **x + σ**, 95.5% of all results lie between **x - 2σ** and **x + 2σ**, and 99.7% of all results lie between **x - 3σ** and **x + 3σ**. Thus, if you make a very large number of measurements, you have "68% confidence" that a future measurement will fall within **± σ** of the average. In this sense, the standard deviation can be regarded as the uncertainty in one measurement, for example, the next measurement you will make.

A more useful quantity is the "uncertainty in the average" of all **N** measurements. What this means is, if you were to take many "sets" of **N** measurements and find the average for each, how close together would the averages be? The answer is that the averages would be distributed in a bell shaped curve of their own, but the standard deviation of the averages would be smaller than the standard deviation of the data in any one "set". This is why it's better to take a large number of measurements and find the average than to just take a single measurement! It turns out that the uncertainty in the average is √**N** times smaller than the standard deviation.

Rule #11: *The uncertainty in the average of a large number, N, of measurements may be taken as the standard deviation of the data divided by √N.*