UNCERTAINTY

an introductory introduction

In science, it is important not only to know a value, but to know how precisely that value is known. One way of expressing this that you might be familiar with is *significant digits*. A more general and exact way is by use of *uncertainty*, denoted with the symbol "±" (read "plus or minus"). For instance[1] the value of the constant of gravitation **G** is:

$$\mathbf{G = (6.673 \pm 0.010) \times 10^{-11}\ m^3/kg \cdot s^2}$$

This means roughly that the numerical value of **G** is believed to be between $\mathbf{6.663 \times 10^{-11}}$ and $\mathbf{6.683 \times 10^{-11}}$. We can also talk about a *percent uncertainty*, which is the uncertainty divided by the base value (times **100%**). For **G**, the percent uncertainty is equal to **0.010 / 6.673 = 0.0015 = 0.15%**, and so it can also be written as:

$$\mathbf{G = 6.673 \times 10^{-11}\ m^3/kg \cdot s^2\ (\pm 0.15\%)}$$

Generally, percent uncertainty gives you a better idea of how well a number is known than regular uncertainty. Depending on the experiment, you shouldn't expect to get much lower than **±5%** in the lab, and **±40%** is not unheard of. Uncertainties much less than **±1%** (as for **G** above) usually only come from years of observations.

So it's a good idea to express your final uncertainties as percent uncertainties, to get an idea of how precise the experiment was. However, for some kinds of values, percent uncertainties are not meaningful. These numbers should always have plain uncertainty, never percent uncertainty. These are:

1. Dates and times, like **9:22:15AM**. (However, *durations* of time are okay, such as **10.3 hours (±16%)**.)
2. Positions, such as X-Y coordinates, latitude-longitude, and Right Ascension-declination.
3. Relative scales, such as temperatures in Celsius or Fahrenheit.
4. Logarithmic scales, such as stellar magnitudes and decibel levels.

Because uncertainty is useful in some situations, and percent uncertainty is useful in other situations, it is very valuable to be able to convert between them. So make sure that you are comfortable going back and forth.

Measurement Uncertainty

There is some uncertainty inherent in every measuring device. For something like a ruler, the inherent uncertainty is roughly equal to the smallest mark. For instance, if the smallest units marked on the ruler are centimeters, then you are measuring "to the nearest centimeter", and your measurements will be **±1cm**. Digital watches will typically have an uncertainty of **±1 sec**.

In addition to the uncertainty inherent in the device, every measurement carries some uncertainty based on how difficult it is to determine the exact value. For example, suppose you intend to record the time at which the Sun sets. You may measure the sunset as occurring at **5:48:11PM**, but because it's so hard to tell exactly, it could just as easily have happened two minutes earlier or later. You can express this as **5:48:11PM (±2 min)**. It is tempting to be overly conservative when estimating measurement uncertainty, but accuracy is more important than modesty. Uncertainty values that are too high are equally as bad as values that are too low.

In general, the numbers you list for your measurements should not appear to be more precise than the uncertainty implies. For instance, the above observation should just be listed as **5:48PM (±2 min)**, because the **11 seconds** is a bit less than the uncertainty. Similarly, you would not write the length of the year as $\mathbf{(3.1602398 \pm 0.02) \times 10^7}$ **sec**, but rather $\mathbf{(3.16 \pm 0.02) \times 10^7}$ **sec**.

---

1 National Institute of Standards and Technology, `http://physics.nist.gov/cgi-bin/cuu/Value?bg`

Standard Deviation

If you have a set of data which are supposed to be measurements of the same thing, you can calculate the statistical *standard deviation*, which is a measurement of the spread in the data set. Standard deviation is denoted by the Greek letter σ (sigma). For instance, consider the data set **{25, 28, 31, 26, 28}**. The average of this data set is **27.6** and the standard deviation is $\boldsymbol{\sigma}$ **= 2.3**. So the five values in the data set may be reduced into a single value: **27.6±2.3**, or **27.6 (±8%)**. It is possible but tedious to calculate σ by hand.[2] Fortunately, you should have access to a calculator that will do it for you.

To determine the average and standard deviation of a data set using Windows Calculator:

1. Under the **View** menu, make sure **Scientific** is selected.
2. Pick the **Sta** button. The **Statistics Box** will appear. Pick the **RET** button to return to the calculator.
3. Enter each of the data values, followed by the **Dat** button.
4. Once all the data are entered, get the average by picking the **Ave** button, and the standard deviation by picking the **s** button.

Most scientific calculators will perform similar functions. For the TI-89, entering data is described on p. 242 of the manual and getting statistical information is described on p. 247-248. You can also enter it on the command line using the form:

```
stddev({25,28,31,26,28})
```

Calculating with Uncertainty – Error Propagation

When performing calculations on numbers with uncertainty, you must also perform calculations on the uncertainty itself. The mathematics of calculating with uncertainty is called Error Propagation.[3] For some kinds of calculations, it's easier to work with normal uncertainty, but for some, it's easier to work with percent uncertainty.

When *adding or subtracting two values*, each with its own uncertainty, add the uncertainty. For example, if the orbit of Mercury is **(5.79±1.19)×10^{7} km** in radius, and the orbit of Venus is **(1.082±0.007)×10^{8} km** in radius, then the distance between their orbits is:

$$(1.082\times10^8 - 5.79\times10^7)\text{ km} \pm (0.007\times10^8 + 1.19\times10^7)\text{ km}$$
$$= 5.03\times10^7\text{ km} \pm 1.26\times10^7\text{ km}$$
$$= (5.03 \pm 1.26)\times10^7\text{ km}.$$

When *multiplying by an exact number*, percent uncertainty is maintained. For example, suppose you know the radius of the Sun is **R = 7.2×10^{5} km (±9%)**. Then the circumference of the Sun, given by **2πR**, will be **4.5×10^{6} km (±9%)**. Converting units also leaves percent uncertainty intact. In the previous example, the circumference could also be expressed as **4.5×10^{11} cm (±9%)** or **2.8×10^{6} miles (±9%)**.

When *multiplying or dividing two values*, each with its own uncertainty, add the percent uncertainty. For example, suppose a comet moves **6×10^{7} km (±21%)** over the course of **19.5 days (±2%)**. Then, using **speed = distance / time**, we can get the comet's speed:

$$6\times10^7\text{ km} / 19.5\text{ days}$$
$$= 35.6\text{ km/s } (\pm23\%).$$

When you *raise a value to a power*, multiply the percent uncertainty by that power. For example, Kepler's Third Law states that for the orbital period **P** in years and the separation **a** in AUs, **P = a$^{3/2}$**. So if **a = 4.1 AU (±10%)**, then **P = 8.3 years (±15%)**. Remember that taking the square root is the same as raising to the **1/2** power, and taking the **n**th root is the same as raising to the **1/n** power.

---

2 If you really want to, though, you can. The formula is easy to find, and the following page has an example: `http://news.morningstar.com/news/ms/Investing101/riskybusinesstwo.html`

3 Chapter 5 of the following manual gives more detail on the formulae for Error Propagation, and some examples: `http://www.rit.edu/~vwlsps/uncertainties/Uncertaintiespart2.html`

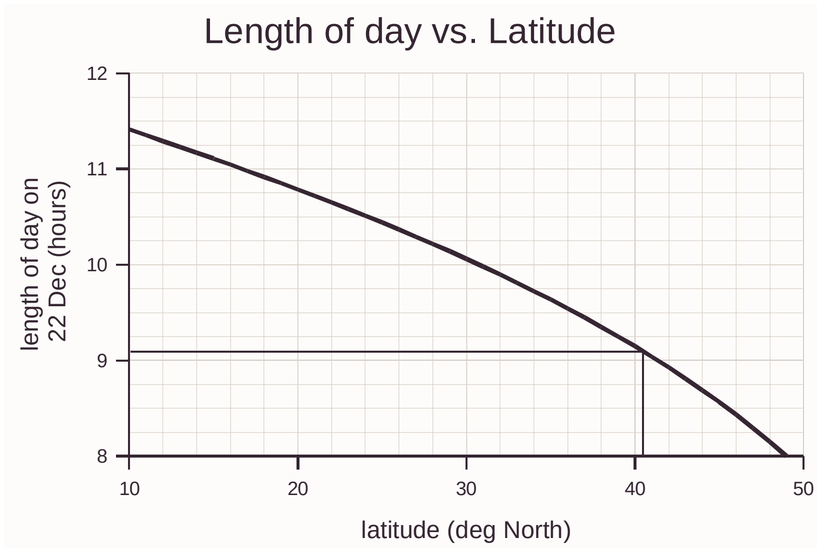
When performing a calculation that does not fit one of these, you have to do a little more math. As an example, consider the distance modulus equation

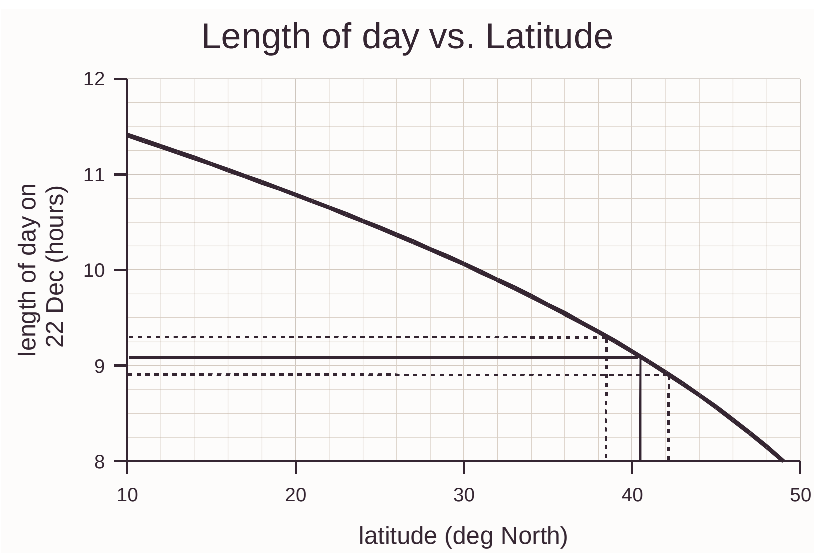
$$\mathbf{DM = 5 \log (d / 10)},$$

where **d** is in parsecs. Suppose you have **d = (206±10)pc**, and you wish to determine the distance modulus **DM**. Since **log** is not covered by any of the situations in the previous paragraphs, you must perform the calculation three times – once on **206pc**, once on **(206+10)pc**, and once on **(206-10)pc**. The base value of **d = 206pc** results in **5 log(20.6) = 6.57**. This will be your base value for **DM**. Putting the other two values in the equation results in **6.67** and **6.46**. Since these are **+0.10** and **-0.11** from the base value, a reasonable average would be **±0.11**. (It's okay if the two values don't match up; just pick a rough average.) So the result is **DM = 6.57±0.11**.

Reading Uncertainty from a Graph

Sometimes you will need to determine one value from another by reading from a graph instead of using an equation. For example, the graph at the right relates the length of the day (from sunrise to sunset) on the Winter Solstice to the observer's latitude in the Northern hemisphere. This lets you use latitude to determine the length of day, and vice versa. Suppose that you measured the length of day as **9 hours, 6 min**, that is, **9.1 hours**. To determine the latitude from this, find **9.1 hours** on the y-axis of the graph, draw a horizontal line over to the curve, and then draw a vertical line down to the x-axis, as shown on the graph. In this case, **9.1 hours** corresponds to **40.4°N** latitude.





Now consider the same problem taking uncertainty into account. Suppose the length of day measurement is **9.1±0.2 hours**. In order to translate this uncertainty in length of day into an uncertainty in latitude, draw two more lines on the graph, at **9.1+0.2 hours** and **9.1-0.2 hours**. These correspond to x-axis values of **42.2°N** and **38.5°N** latitude, as shown in the second graph with dashed lines. So you would conclude that a length of day of **9.1±0.2 hours** corresponds to **(40.4±1.8)°N** latitude.