Analysis of Uncertainties   **Fall 2014**

Lecture 1: 20 October 2014

Lecturer: Prof. Kate Scholberg   *Notes by: Douglas Davis*

1.1   Introduction

Think about the value for Newton's gravitational constant G; the current accepted value in literature is something around (call it G_{m1}, this doesn't have to be correct we're just making a point).

$$G_{m1} = (6.674215 \pm 0.000092) \times 10^{11} \ \mathrm{m}^3 \ \mathrm{kg}^{-1} \ \mathrm{s}^{-2}.$$

Now imagine we perform some experiments where we measure G. Imagine we measure a a new value, call it G_{m2},

$$G_{m2} = (6.674400 \pm 0.000010) \times 10^{11} \ \mathrm{m}^3 \ \mathrm{kg}^{-1} \ \mathrm{s}^{-2}.$$

And we measure another value, G_{m3},

$$G_{m3} = (6.6741 \pm 0.0221) \times 10^{11} \ \mathrm{m}^3 \ \mathrm{kg}^{-1} \ \mathrm{s}^{-2}.$$

If we look closely at the values and the uncertainties, we see that the value G_{m2} is in disagreement with G_{m1}, because the values do not lie within the other value's uncertainty. This means some further investigation is needed! If we look at G_{m3}, we see that it is in agreement with both G_{m1} and G_{m1}, but the uncertainty is so large that this measurement really did nothing for the advancement of science. This can be seen graphically in Figure 1.1. The point of this example is to show that uncertainties seriously matter. When interpreting
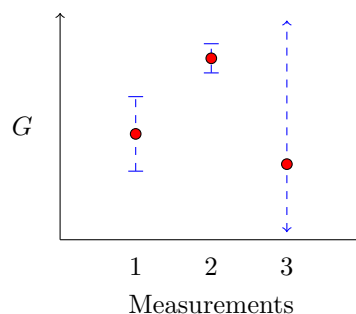

Figure 1.1: Three measurements for G, the values are red bullets, and the uncertainties are shown with blue dashed error bars

results and planning new experiments, a scientist must always be thinking about uncertainties. Why build an experiment if the expected uncertainty in the measurement will yield nothing better than what is already known? Or, what does this new measurement mean that is it so much more constrained than the previous measurement? Is this method so great we should expand it elsewhere, and why is it so much better?

1.2   Important Definitions

Our first two important definitions are *statistical* and *systematic uncertainties*. A *statistical uncertainty* is a random uncertainty that arises simply from random fluctuations in the data. As the sample size of the data approaches infinity ($N \to \infty$), the statistical uncertainty approaches 0. A *systematic uncertainty* comes from a lack of knowledge about the experimental apparatus/measurement. An example would be the possibly of an inexact measurement from an imperfect ruler. Systematic uncertainties can also have some randomness, if an apparatus is affected by temperature change, the random fluctuations of temperature in the room cause systematic uncertainties because of the effect on the apparatus. To quote the systematic and statistical uncertainties, results are often written in the form:

$$x \pm \Delta x \,(\text{sys.}) \pm \Delta x \,(\text{stat.}), \tag{1.1}$$

with the two values separate. As the amount of data increases, the statistical uncertainty always decreases. The same cannot be said for systematic. In theory, studying systematic uncertainties more thoroughly may lead to a decrease, but it is possible that a better understanding of the systematics may cause an increase in the uncertainty. Our next definition is a *sample*. The *sample* is the collection of measurements. If we want to measure some value x, then the sample is simply:

$$x_1, \, x_2, \, x_3, \, ... \, , \, x_N,$$

where x_i is just the ith measurement. An infinite number of measurements would yield the sample's *parent distribution*, which we denote with $P(x)$. Figure 1.2 shows an example of a histogrammed sample with it's parent distribution. As $N \to \infty$, the shape of the histogram would approach the shape of $P(x)$.
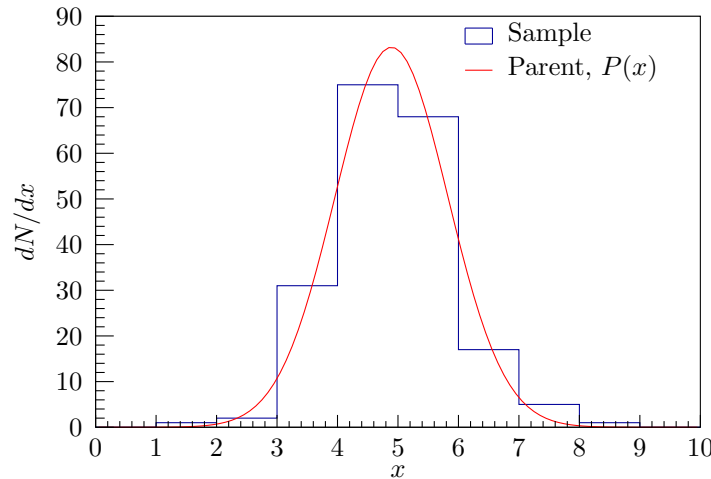

Figure 1.2: A sample distribution in histogram form overlayed by its parent distribution.

Now we will define some characteristics of the sample and the parent distribution. First, the *mean of the sample*, which we denote with $\overline{x}$:

$$\overline{x} = \frac{1}{N} \sum_i^N x_i. \tag{1.2}$$

The *mean of the parent distribution*, denoted by μ:

$$\mu = \lim_{N \to \infty} \frac{1}{N} \sum_i^N x_i. \tag{1.3}$$

To quantify the spread we may first think to use the *deviation*, denoted by d,

$$d_i = x_i - \mu, \tag{1.4}$$

and take the limit:

$$\lim_{N \to \infty} \frac{1}{N} \sum_i^N d_i = \lim_{N \to \infty} \frac{1}{N} \sum_i^N (x_i - \mu), \tag{1.5}$$

but because d_i can be positive or negative, this limit tends to zero. We may also think to take the absolute value of the deviation, $|d_i|$, but this math turns out to be pretty bad. Instead we go with something called the *variance*, which we denote with σ^2:

$$\sigma^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i^N (x_i - \mu)^2. \tag{1.6}$$

Which can also be written as:

$$\sigma^2 = \lim_{N \to \infty} \left[\frac{1}{N} \sum_i^N x_i^2 \right] - \mu^2 \tag{1.7}$$

In real world experiments, we probably do not know the parent mean, so we take $\mu \to \overline{x}$. Now we define the *sample variance*, denoted by s^2:

$$s^2 = \frac{1}{N-1} \sum_i^N (x_i - \overline{x})^2 \tag{1.8}$$

Notice the difference between the fraction in equation 1.6 and equation 1.8, we change from $1/N$ to $1/(N+1)$. The reason for this can be worked out mathematically (ref), but we will just explain the following reason. Imagine taking only 1 measurement ($N = 1$). Equation 1.6 would yield a variance of only 1. This does not make sense for a single measurement, the outcome using equation 1.8 would be an undefined variance, which does make sense. There is no variance in a single measurement. The sample variance defined in equation 1.8 is known as an *unbiased estimator*. Closing side note: for practical purposes many scientists use the notation of σ^2 for the sample variance and μ for the sample mean.

1.3   Propagation of uncertainties

Let's suppose the value we are attempting to measure, x, is a function of u:

$$x = f(u) = u^2 \tag{1.9}$$

We know that our measurement, x, will have some uncertainty (which we will denote as δx or σ_x), and our independent variable must also have some uncertainty: δu, σ_u. We can (to first order, assuming δx is small) write:

$$\delta x = f(u + \delta u) - f(u). \tag{1.10}$$

We can then write:

$$\frac{\delta x}{\delta u} = \frac{f(u + \delta u) - f(u)}{\delta u} = \frac{df}{du} \rightarrow \delta x = \left| \frac{df}{du} \right| \delta u, \tag{1.11}$$

or:

$$\sigma_x = \left| \frac{df}{du} \right| \sigma_u \tag{1.12}$$

Therefore, the measurement related to equation 1.9 would be of the form:

$$\text{measured value} = x \pm |2u|\sigma_u \tag{1.13}$$

An example of this would be the area of a square with side length u, and an uncertainty in the ruler used to measure the side which is σ_u.

Now let's expand this to an arbitrary number of independent variables. Let's take x in the form of f again:

$$x = f(u, v, w, ...), \tag{1.14}$$

it follows that:

$$\overline{x} = f(\overline{u}, \overline{v}, \overline{w}, ...) \qquad x_i = f(u_i, v_i, w_i, ...) \tag{1.15}$$

We can write:

$$x_i - \overline{x} \cong (u_i - \overline{u}) \left(\frac{\partial x}{\partial u} \right)_{v,w,...} + (v_i - \overline{v}) \left(\frac{\partial x}{\partial v} \right)_{u,w,...} + ... \tag{1.16}$$

Now recall equation 1.6, we can write:

$$\sigma_x^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i^N \left[(u_i - \overline{u})^2 \left(\frac{\partial x}{\partial u} \right)^2 + (v_i - \overline{v})^2 \left(\frac{\partial x}{\partial v} \right)^2 + ... + \underbrace{2(u_i - \overline{u})(v_i - \overline{v}) \left(\frac{\partial x}{\partial u} \right) \left(\frac{\partial x}{\partial v} \right) + ...}_{\text{cross terms from squaring eq. 1.16}} \right] \tag{1.17}$$

Now using our standard variance equation with the form:

$$\sigma_u^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i^N (u_i - \overline{u})^2, \tag{1.18}$$

and now introducing the covariance between u and v:

$$\sigma_{uv}^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i^N (u_i - \overline{u})(v_i - \overline{v}). \tag{1.19}$$

We can now rewrite equation 1.17 in the form:

$$\sigma_x^2 \cong \sigma_u^2 \left(\frac{\partial x}{\partial u} \right) + \sigma_v^2 \left(\frac{\partial x}{\partial v} \right) + ... + 2\sigma_{uv}^2 \left(\frac{\partial x}{\partial u} \right) \left(\frac{\partial x}{\partial v} \right) + ... \tag{1.20}$$

If u and v are independent of eachother, in the large N limit, $\sigma_{uv}^2 \rightarrow 0$ and all of the "cross-terms" vanish, yielding a variance of:

$$\sigma_x^2 \cong \sigma_u^2 \left(\frac{\partial x}{\partial u} \right) + \sigma_v^2 \left(\frac{\partial x}{\partial v} \right) + ... \tag{1.21}$$