Error Analysis in PHYS180

Jordan C. Hanson

March 9, 2018

Abstract

Scientific measurements cannot be quoted without an assessment of the precision. We quantify this precision in the form of error analysis. Rather than originating from *human error*, *statistical errors* arrise from the intrinsic uncertainty encounted in measuring quantities with real instruments. Further, errors *propagate* through calculations.

1 Instrumental Precision

The instrumental precision of a piece of data is determined by the smallest division on the instrument. For example, suppose a length is measured with a meter stick, and the smallest divisions on the meter stick are 1 millimeter. If the length measured is 31.0 cm, then that result must be stated as 31.0 ± 0.1 cm.

2 Accuracy versus Precision

Accuracy and precision are not the same concept. Accuracy is how far the measured result is from the true value, whereas precision compares the statistical error of a measurement to the measured value itself. A measurement of 31.0 ± 0.1 cm is precise, because 0.1 cm is small compared to 31.0 cm. It is not *accurate* if the true length is 29.0 cm. Why? Because the numbers 31.0 cm and 29.0 cm are separated by $(31.0 - 29.0)/0.1 = 20$ factors of the error.

3 Statistical or Random Error

Suppose a measurement is made with a dial, where a needle is pointing to a number, but vibrating. When the dial is read one moment, it reads 101.3 kPa, and in another moment, 101.2 kPa. Which pressure measurement of the atmosphere is correct? It turns out that upon repeated measurements, the *distribution* of measurements resembles Fig. 1.

The true pressure is most likely 101.3 kPa. The dial needle is moving, however, making identification of the true pressure impossible. The true value is ascertained from the statistical distribution. The *mean* or average of the distribution is 101.3 kPa, and the *standard deviation* is about 0.2 kPa. Equation 1 represents the standard deviation procedure. First, the mean $\bar{x}$ is calculated for the collection of measurements x_i. Next, the quantity $(x_i - \bar{x})^2$ is computed for each x_i. Next, the sum of these quantities is found, dividing by N[1], where N is the number of measurements.

---

[1] Technically, we should divide by $N - 1$ since the mean has already accounted for one degree of freedom, but in practice where N is a large number, this rarely matters.
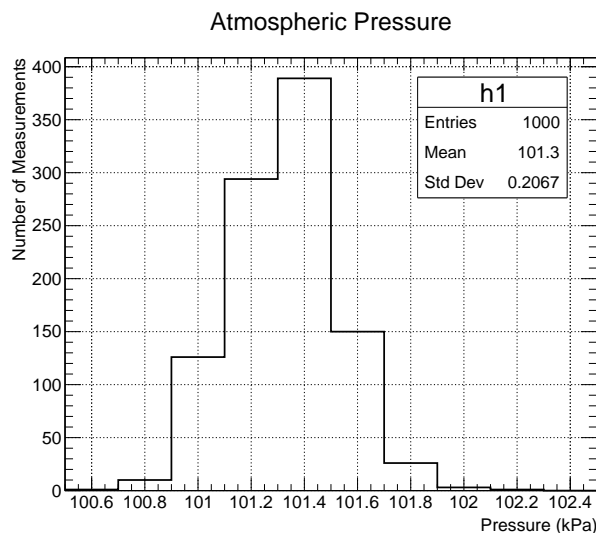
Figure 1: An example of a distribution of measurements from one instrument: a pressure gauge.

L_f (kJ/kg)
310
389
267
385
319
402
322
341

Table 1: Some measurements of the laten heat of fusion. The average is 340 kJ/kg, and the standard deviation is 50 kJ/kg.

This result is called the *variance*, σ^2. Finally the square root of the variance is taken to find the standard deviation: $\sqrt{\sigma^2} = \sigma$.

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

In class, we made measurements of the **latent heat of fusion** L_f of water. Those measurements, along with a few more, are listed in Tab. 1. The average latent heat is 340 ± 50 kJ/kg. We quote the mean, plus or minus the standard deviation as a sign of the precision of our measurement.

4   Analysis of Data Subject to Random Error

4.1   Addition and Subtraction of variables with errors

It turns out that computing the mean and standard deviation of a list of results is a different statistical question than assessing the error in a measurement when that measurement is a combination of other measurements. We measured the latent heat of water by measuring the initial

masses and initial temperatures of hot water and ice, and a final temperature of the mixture. The relevant equation is

$$L_{f,H2O} = c_{H2O} \left(\frac{m_{H20}}{m_{ice}} (T_{i,H20} - T_f) - T_f \right) \tag{2}$$

The coefficient c_{H20} is the specific heat of water, m_{ice} is the ice mass, m_{H20} is the water mass, $T_{i,H20}$ is the temperature of the hot water, and T_f is the final mixture temperature. In Tab. 1, the mean and the standard deviation represent the average value of the measurements obtained by the class, and the scatter or spread in those measurements. This is different than assessing the error in *each measurement of L_f*. We saw that the mass measurements were highly precise, maybe 0.5 grams of error. Let's ignore errors associated with mass.

What happens to the error in Eq. 2 if we begin to think about errors in temperature? Notice that two temperatures are subtracted, then multiplied by a mass ratio, and then another temperature is subtracted. As an example, let's suppose a group used hot water with $T_{i,H20} = 80 \pm 2$ °C, and equal masses of ice and water, and observes a final temperature of $T_f \pm 2$ °C. Working in the units of calories, grams, and degrees Celsius, Eq. 2 reduces to

$$L_{f,H20} = (T_{i,H20} - 2T_f) \quad \frac{cal}{g} \tag{3}$$

Notice that the group decided that their measurements of temperature were accurate to only about two degrees. There are two operations happening to temperature in Eq. 3. Multiplication by two, and subtracting two temperatures. If a variable x has an error σ_x associated with it, and it is multiplied by two, the result is

$$2(x \pm \sigma_x) = 2x \pm 2\sigma \tag{4}$$

In other words, the error doubles. In general, for some constant k,

$$k(x \pm \sigma_x) = kx \pm k\sigma_x \tag{5}$$

Regarding addition and subtraction of variables $x \pm \sigma_x$ and $y \pm \sigma_y$, it turns out that

$$(x \pm \sigma_x) + (y \pm \sigma_y) = (x + y) \pm \sqrt{\sigma_x^2 + \sigma_y^2} \tag{6}$$

$$(x \pm \sigma_x) - (y \pm \sigma_y) = (x - y) \pm \sqrt{\sigma_x^2 + \sigma_y^2} \tag{7}$$

Applying these rules to Eq. 3 to account for errors yields

$$L_{f,H20} = (T_{i,H20} - 2T_f) \pm \sqrt{\sigma_{i,H20}^2 + (2\sigma_f)^2} \quad \frac{cal}{g} \tag{8}$$

The group used two degrees for the error in the initial and final temperatures in Eq. 8, and 80 degrees and 0 degrees were the observations. **Plugging this information into Eq. 8 yields** 80 ± 4.5 calories per gram. The expected value is 80 calorie per gram, and the error is $4/80 = 5\%$ of the result. Dividing the *standard deviation by the result is called the fractional error*. In the next lesson on error analysis, we will learn to account for multiplication and division of variables through the use of fractional errors.

4.2 Multiplication and Division of variables with errors

If two variables are multiplied or divided, the *fractional errors* are combined rather than the *absolute errors.* Suppose we are trying to quantify the momentum of a system with mass m and speed v:

$$p = mv \tag{9}$$

We measure the mass: $m \pm \sigma_m$, and the speed: $v \pm \sigma_v$. The *fractional error* in the momentum, speed, and mass are $\frac{\sigma_p}{p}$, $\frac{\sigma_v}{v}$, and $\frac{\sigma_m}{m}$ respectively. The relationship between the fractional error in momentum and the other two is

$$\left(\frac{\sigma_p}{p}\right)^2 = \left(\frac{\sigma_m}{m}\right)^2 + \left(\frac{\sigma_v}{v}\right)^2 \tag{10}$$

Notice that the fractional errors of the mass and speed are added in quadrature (summing squares) to produce the frational error of the momentum. Thus, to quote a measurement of momentum, the result is:

$$p = (mv) \pm (mv)\sqrt{\left(\frac{\sigma_m}{m}\right)^2 + \left(\frac{\sigma_v}{v}\right)^2} \tag{11}$$

5 Example with Addition and Subtraction, Multiplication and Division

Adiabatic work. The work done by an adiabatic process is given by

$$W = \frac{1}{1 - \gamma}\left(p_f V_f - p_i V_i\right) \tag{12}$$

Let's assume we are working with a monatomic gas with only three degrees of freedom. This makes $\gamma = c_P/c_V = (c_V + R)/c_V = (3/2 + 1)/(3/2) = 5/3$. Suppose a process takes a system from a state $(1.0 \pm 0.05$ L$,2.0 \pm 0.1$ atm$)$ to $(2.0 \pm 0.1$ L$,0.5 \pm 0.1$ atm$)$. What is the work done, accounting for random statistical errors?

1. First, let's compute the factor in front of Eq. 12 on the right side, since it just scales everything and doesn't contribute to the error propagation.

$$\frac{1}{1 - \gamma} = \frac{1}{1 - \frac{5}{3}} = \frac{1}{-\frac{2}{3}} = -\frac{3}{2} \tag{13}$$

Even though this pre-factor is negative, the work done is positive, because the area under the pV curve will be positive.

2. We can see that the errors associated with the pressures and volumes with combine following the multiplication rule: fractional errors combine in quadrature. For the error in the each state, we have

$$\frac{\sigma_{pV}}{pV} = \sqrt{\left(\frac{\sigma_p}{p}\right)^2 + \left(\frac{\sigma_V}{V}\right)^2} \tag{14}$$

3. Solving for σ_{pV}, and plugging in numbers for each state with units of liters $\times$ atmospheres:

$$\sigma_{pV} = (pV)\sqrt{\left(\frac{\sigma_p}{p}\right)^2 + \left(\frac{\sigma_V}{V}\right)^2} \tag{15}$$

$$\sigma_i = (1 \cdot 2)\sqrt{\left(\frac{0.1}{2.0}\right)^2 + \left(\frac{0.05}{1.0}\right)^2} = 0.1414 \ \ L \cdot atm \tag{16}$$

$$\sigma_f = (1 \cdot 2)\sqrt{\left(\frac{0.1}{0.5}\right)^2 + \left(\frac{0.1}{2.0}\right)^2} = 0.4123 \ \ L \cdot atm \tag{17}$$

4. Now we have the errors for the initial and final states, $p_i V_i$ and $p_f V_f$. These states are subtracted, then multiplied by $-\frac{3}{2}$. The multiplication by this factor just scales the error, so let's save that for last. Absolute errors combine in quadrature when we subtract or add them:

$$\sigma = \sqrt{\sigma_i^2 + \sigma_f^2} = \sqrt{(0.1414)^2 + (0.4123)^2} = 0.4358 \ \ L \cdot atm \tag{18}$$

5. Now we multiply by $\frac{3}{2}$ (error can be positive or negative, so it doesn't really matter that the pre-factor is $-\frac{3}{2}$):

$$\sigma \times \frac{3}{2} = 0.6537 \ \ L \cdot atm \tag{19}$$

6. Quoting the results:

$$W \pm \sigma = 1.5 \pm 0.65 \ \ L \cdot atm = 150 \pm 65 \ \ J \tag{20}$$