

# 1 Precision Estimates

## Introduction

Physics is sometimes called an ‘exact’ science. But this does not mean that the numerical uncertainties in measured quantities, or even in the laws of physics, are vanishingly small. Rather, it reflects the empirical fact that with care one often can make measurements whose uncertainties, or ‘errors’, are smaller than appear important.

In the Ph101 laboratory you will have the opportunity to make several measurements, and likely in later years you will have the need to evaluate the merits of measurements made by others. Typically a good understanding of the precision of the measurements is at least as important as the measured value itself.

This Laboratory will use three types of measurements to illustrate procedures for estimating the precision (the optimist’s view) or the error (the pessimist’s view) of measurements.

1. The density of aluminum. This is an example of a quantity whose value could be well determined with sufficient effort, but the limitations of the instruments lead to an uncertainty in the measurement that is sometimes called measurement error.
2. The thickness of a dollar bill. Although the thickness of dollar bills may vary, the thickness is so small that measurement uncertainty is significant here also.
3. Human reaction time. In this case the instruments are more precise than the variation in repeated measurements of the quantity, which has an intrinsic uncertainty.

### 1.1 The Density of Aluminum

The Handbook of Chemistry and Physics (reference copy available in the Ph101 Laboratory) lists the density of aluminum as 2.69 gm/cm<sup>3</sup> on one page and as 2.70 gm/cm<sup>3</sup> on another. Possible interpretations are that the density of various batches of aluminum varies by 0.01 gm/cm<sup>3</sup> due to impurities, or that 0.01 gm/cm<sup>3</sup> is the uncertainty in the measurement of the density.

Make your own determination of the density,  $\rho$ , of aluminum by measuring the volume and mass of a sample block:

$$\rho = \frac{M}{V} = \frac{M}{L \cdot W \cdot H}, \quad (1)$$

where  $L$ ,  $W$  and  $H$  are the length, width and height, respectively, of the block. For each of the four measured quantities you should make a simple estimate of the uncertainty in that measurement. Labeling the uncertainty on  $M$  as  $\sigma_M$ , etc., then the error on the density follows from the equation (16) of the Appendix as

$$\frac{\sigma_\rho}{\rho} = \sqrt{\left(\frac{\sigma_M}{M}\right)^2 + \left(\frac{\sigma_L}{L}\right)^2 + \left(\frac{\sigma_W}{W}\right)^2 + \left(\frac{\sigma_H}{H}\right)^2} \quad (2)$$

*Measure the mass with either the digital or analog balances on the tables near the center of the laboratory. A typical estimate of the uncertainty of a digital device would be the*

value of the least count – although this assumes the engineers have designed the device to this standard. For an analog device such as a balance or ruler a typical estimate of the uncertainty would be about  $1/3$  of the smallest gradation. See sec. 1.5.5 for a technical comment on this.

However, if repeated measurements of a quantity give varying results, the uncertainty is larger than the above estimate; the quantity being measured is uncertain. In this case you should characterize the spread of your data by a ‘standard deviation’ as discussed in the Appendix. Start with the simple approach of sec. 1.5.2 which is based on 3 measurements.

To measure lengths you can use meter sticks, rulers, vernier calipers or micrometers, in order of decreasing range of lengths that can be measured. To minimize the measurement uncertainty, use the device with the smallest gradations that also spans the length you wish to measure. For example, use the micrometer to measure the smallest dimension of the aluminum block.

Both the vernier caliper and the micrometer utilize the vernier technique to improve the accuracy of the measurement. Be sure to master the use of a vernier during this laboratory.

The aluminum blocks may have a small conical hole drilled in them. If so, subtract the volume of the hole from the volume of the block before applying eq. (1). You may omit an error analysis of the volume of the cone, but compare  $V_{\text{hole}}/V_{\text{block}}$  to  $\sigma_{\rho}/\rho$  to estimate the importance of the volume of the cone for the density measurement.

Is your measurement of the density  $\rho$  of aluminum within one or two times your error estimate  $\sigma_{\rho}$  of the book value 2.70? If not, you have probably made a mistake somewhere. Check your arithmetic; recheck your measurements if necessary. One of the advantages of performing the error analysis is the confidence it gives that your measurements are self consistent when both measurements and errors appear reasonable; or conversely the clue it gives you if the measured value varies from expectations by more than the error estimate.

**Each person should take his or her own measurements on the aluminum block and record them and the appropriate data analysis in their lab notebook before proceeding.**

## 1.2 Thickness of a Dollar Bill

In the following part of the lab you can use a dollar bill to estimate your reaction time. First, measure the thickness of a dollar bill, and estimate the error on your measurement.

*Since a dollar bill is so thin, it's better to measure the thickness of a stack of  $n$  dollar bills, then divide by  $n$ . In case you don't have a thick wad of bills, fold a single bill several times. Note that paper is highly compressible, so only the uncompressed thickness of a dollar bill is well defined.*

If  $t$  is the thickness of a bill and you fold it until there are  $n$  layers to measure, you will measure thickness  $T = nt$ . If  $\sigma_T$  is your estimate of the uncertainty on the measurement of total thickness  $T$ , then the corresponding uncertainty on  $t = T/n$  is  $\sigma_t = \sigma_T/n$  since there is no uncertainty in the number  $n$ .

How many dollar bills would there be in a stack high enough to reach the moon, whose distance is about 400,000 km? What is the uncertainty on your estimate?

## 1.3 Human Reaction Time

### 1.3.1 Catching a Dollar Bill

A well-known challenge is to catch a dollar bill dropped by your partner, holding your thumb and finger in line with the picture of George W. before your partner lets the bill go. Human reaction time is such that very few people can catch the bill (without anticipating the release by some clue). From the expression for the distance travelled by a falling object from rest,

$$y = \frac{1}{2}gt^2, \quad (3)$$

we infer that the human reaction time  $t_{\text{react}}$  is longer than  $\sqrt{2y/g}$  where  $y$  is the half-length of a dollar bill and  $g$  is the acceleration due to gravity.

Try the experiment, and record the numerical value of eq. (3) as a first estimate of  $t_{\text{react}}$  in your lab book.

### 1.3.2 Catching a Meter Stick

The principle of sec. 1.3.1 can be used to make a measurement of your reaction time.

*Have your partner hold a meter stick from its upper end, using the plate attached to the vertical rod on your lab bench to steady his/her hand. Place your thumb and finger in line with some gradation near the bottom of the stick. Catch the stick when it is dropped and record the distance through which it fell before you caught it. Use the analysis of sec. 1.3.1 to deduce your reaction time.*

*Repeat your measurement at least three times. Each person in your group should collect data on his/her own reaction time. Report the average time of your results as your reaction time. Estimate the error on your measurement according to the simple prescription of sec. 1.5.2: in the case of three measurements, the error estimate is one half the difference between the highest and lowest measurements. Either convert each measurement of distance to a time and analyze the time spread; or analyze the spread of distances and use eqs. (3) and (16) to propagate the error according to*

$$\frac{\sigma_t}{t} = \frac{1}{2} \frac{\sigma_y}{y}. \quad (4)$$

### 1.3.3 Electronic Timing

As the final measurement of your reaction time you can use the timing capabilities of the PC by the lab bench. Your partner will start the timer via a push-contact and you will stop it as quickly as possible via a second push contact. Repeat the measurement 25 to reduce the uncertainty in your result from fluctuations.

*If the PC is in Windows then Exit. The dos directory for the timer software is c:\timer. Start the program by typing pt. From the main menu select Pulse Timing Modes using the arrow keys, then Pulse 1-2, then choose one of the display types such as Times Displayed in Large Digits and press Enter/Return). The timer is now live.*

*Determine which key is the start and which is the stop, and practice taking measurements. The person who starts the timer should avoid giving advance clues as to when he/she will start! A data set is ended by typing Enter on the PC, and you can cycle thru the menus to start a new set if needed.*

*The computer may be set up to emit a sound once the timer is started. You can turn this option on or off via Other Options on the main menu.*

*After a data set has been collected press Enter and then select Display Table of Data. You can view a summary of the results by scrolling down with the arrow key, or jumping to the end with the End key. The average = Mean, standard deviation = SD and standard deviation of the mean = SDOM have been calculated using well-known formulae found in secs. 1.5.4 and 1.5.6. The standard deviation of the mean is the error estimate on your reaction time.*

*Does the new measurement of your reaction time agree within errors with that found in part 1.3.2?*

*If you feel the results are not of sufficient quality, repeat the whole set of measurements (rather than, say, using the computer to edit out data points that you don't like).*

*If you are satisfied with the results, save your timer data to a hard-disk file. From the Data Analysis Options menu select File Options and then Save Data to a File. Enter a unique file name of up to 8 characters and add .dat. Write this file name in your lab notebook!*

*Use Print Table of Data to obtain a copy for your lab notebook.*

*Occasionally the print functions hangs in program pt. To reinitialize, first Save any useful data, then Quit the pt program. Type Install; choose VGA, then Hewlett-Packard Laserjet IIP, then Portrait – Full Page. Accept the default options for the next two screens to exit the Install program. Re-enter program pt and Load Data from Disk if appropriate.*

## 1.4 Histogram of Your Reaction Times

As a final step in the analysis of your reaction time it is useful to make a histogram of the data. This graph can give a quick check as to whether your numerical analysis was reasonable, and could reveal any unusual patterns in the data.

A histogram is readily made by hand on quadrille-ruled paper. But you may prefer to use the StatMost spreadsheet statistical analysis package that runs under Windows on the PC. This will serve as an introduction to some of the PC-based analysis tools that you can use in lab throughout the course.

A fairly detailed procedure for this follows. See how many of the steps you can accomplish without reference to the writeup.

*Quit the timer program from its main menu, and type win to start windows. Then select the Physics 101 program group, and then StatMost. Select File, then New, then Sheet Document to open a new spreadsheet. Again select File and then Import Data. Find the c:\timer\data directory and your .dat file; then Open it.*

*In the window that pops up choose Data Type as Delimited by Commas. Your data should appear in the preview columns A and B. However, the program pt has put header and trailer rows in the file that StatMost cannot use. Select only those rows that contain the timer data, probably beginning with row 11 and not including the final row labelled "EOF", by entering appropriate values in the Row boxes of Import Range. Also restrict the Import Range*

to Columns From 1 to 2. Click on Import to transfer your data into the spreadsheet. Save your spreadsheet now, using the same name as your .dat file but with a .dmd extension.

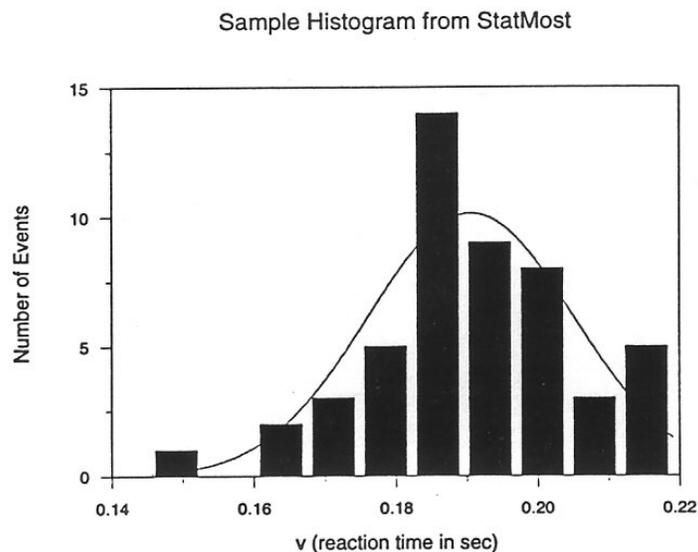


Figure 1: Example of a histogram of measurements of a variable  $v$ . Each measurement is represented by a box of unit height on the plot. An error estimate can be obtained from the histogram by determining the interval in  $v$  that contains the central 2/3 of the measurement.

To produce the histogram select Plot, then 2D Special, then Histogram. In the window that pops up enter your own Title; chose the User Defined option. If you select Draw Normal Curve a bell curve will be fitted to your data and drawn on the plot. Choose Num. of Intervals in the range 5-10; some playing is needed here. Choose the column that contains the data you wish to histogram, probably column B. Click on OK to display the plot; you may have to drag the boundary of (or maximize) the plot window to see its full extent. If the plot looks good, Print it from the File menu and tape it into your lab book. Figure 1 shows a histogram based on reaction-time data.

You can alter the titles and axis labels, or even add additional labels to StatMost plots after they have been created by double clicking on the desired item and using the window that pops up.

Check that your data file still has the same mean and standard deviation as found in the program pt. Click on the spreadsheet window to reactivate it, then select Statistics, then Descriptive Statistics, then General. In the window that pops up choose the column that contains the times (column B), and whatever statistical quantities you wish: Mean, Std Dev and Std Error are useful. Click on OK to generate a window with the summary statistics.

If the histogram of your data doesn't fit reasonably well to the bell curve you may have some poor data points. In such a case consider retaking the data.

## 1.5 Appendix: Error Estimation

While the subject of ‘error analysis’ can become quite elaborate, we wish to emphasize a basic but quite useful strategy, discussed in secs. 1.5.1 and 1.5.2. A more detailed approach is based on the famous bell curve, as sketched in secs. 1.5.3-1.5.5. The important distinction between the error on a single measurement, and the error on the average of many repeated measurements is reviewed in sec. 1.5.6. The subject of ‘propagation of errors’ on measured quantities to the error on a function of those quantities is discussed in sec. 1.5.7.

### 1.5.1 67% Confidence

Whenever we make a measurement of some value  $v$ , we would also like to be able to say that with 2/3 probability the value lies in the interval  $[v - \sigma, v + \sigma]$ . We will call  $\sigma$  the ‘error’ on the measurement. That is, if we repeated the measurement a very large number of times, in about two thirds of those measurements the value would be in the interval stated.

### 1.5.2 A Simple Approach

Repeat any measurement three times. Report the average as the value  $v$ , and the error  $\sigma$  as

$$\sigma = \frac{v_{\max} - v_{\min}}{2}. \quad (5)$$

If you take more than three measurements, you can still implement this procedure with the aid of a ‘histogram’ such as that produced in sec. 1.4. To estimate the error, determine the interval in  $v$  that contains the central 2/3 of the measurements, and report the error as 1/2 the length of this interval.

### 1.5.3 The Bell Curve

In many cases when a measurement is repeated a large number of times the distribution of values follows the bell curve, or gaussian distribution:

$$P(v) = \frac{e^{-(v-\bar{v})^2/2\sigma^2}}{\sqrt{2\pi}\sigma}, \quad (6)$$

where  $P(v)dv$  is the probability that a measurement is made in the interval  $[v, v + dv]$ ,  $\bar{v}$  is the average (or mean) value, and  $\sigma$  is the ‘standard deviation’ which is usually interpreted as the ‘error’. See Figure 2.

The Table lists the confidence that a single measurement from a gaussian distribution falls within various intervals about the mean. If the 200 students in Ph101 each make 50 measurements during this lab, then 10,000 measurements will be taken in all. The Table tells us that if those measurements have purely gaussian ‘errors’, then we expect only one of those measurements to be more than  $4\sigma$  from the mean.

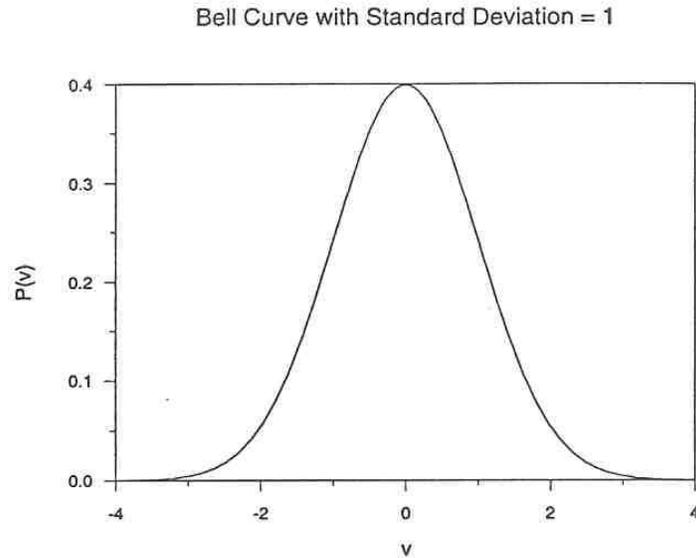


Figure 2: The gaussian probability distribution for mean  $\bar{v} = 0$  and standard deviation  $\sigma = 1$ . The interval  $[\bar{v} - \sigma, \bar{v} + \sigma]$  contains 68% of the distribution.

Table 1: The probability (or confidence) that a measurement of a gaussian-distributed quantity falls in a specified interval about the average.

Interval	Confidence
$\pm\sigma$	68%
$\pm 2\sigma$	95%
$\pm 3\sigma$	99.7%
$\pm 4\sigma$	99.994%

#### 1.5.4 Estimating Errors When Large Numbers of Measurements Are Made

One can make better estimates of errors if the measurements are repeated a larger number of times. If  $n$  measurements are made of some quantity resulting in values  $v_i$ ,  $i = 1, \dots, n$  then the mean is of course

$$\bar{v} = \frac{1}{n} \sum_{i=1}^n v_i, \quad (7)$$

and the standard deviation of the measurements is

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (v_i - \bar{v})^2}. \quad (8)$$

Calculus experts will recognize that the operation  $(1/n) \sum_{i=1}^n$  becomes  $\int P(v)dv$  in the limit of large  $n$ ; then using the gaussian probability distribution given above one verifies that

$$\bar{v} = \langle v \rangle = \int_{-\infty}^{\infty} vP(v)dv, \quad \text{and} \quad \sigma^2 = \langle (v - \bar{v})^2 \rangle = \int_{-\infty}^{\infty} (v - \bar{v})^2 P(v)dv. \quad (9)$$

### 1.5.5 The Error on a Uniformly Distributed Quantity

Not all measurable quantities follow the gaussian distribution. A simple example is a quantity with a uniform distribution, say with values  $v$  equally probable over the interval  $[a, b]$ . It is clear that the average measurement would be  $(a + b)/2$ , but what is the error? If we adopt the simple prescription advocated in secs. 1.1.1 and 1.1.2 we would report the error as  $(b - a)/3$  since  $2/3$  of the time the measurement would fall in an interval  $2(b - a)/3$  long. If instead we use the calculus prescription for  $\sigma$  given in eq. (9) we find that

$$\sigma = \frac{b - a}{\sqrt{12}} = \frac{b - a}{3.46}, \quad (10)$$

which result is often used by experts.

### 1.5.6 The Error on the Mean

Thus far we have considered only the ‘error’ or spread in measured values of some quantity  $v$ . A related but different question is what is the error on the mean value of our measurements,  $\bar{v} = (1/n) \sum v_i$ . The error on the mean  $\bar{v}$  is surely less than the error,  $\sigma$ , on each measurement  $v_i$ . Indeed, the error on the mean is given by

$$\sigma_{\bar{v}} = \frac{\sigma}{\sqrt{n}}, \quad (11)$$

where  $\sigma$  is our estimate of the error on an individual measurement obtained by one of the methods sketched previously.

### 1.5.7 The Error on a Derived Quantity

In many cases we are interested in estimating the error on a quantity  $f$  that is a function of measured quantities  $a, b, \dots, c$ . If we know the functional form  $f = f(a, b, \dots, c)$  we can estimate the error  $\sigma_f$  using some calculus. As a result of our measurements and the corresponding ‘error analysis’ we know the mean values of  $a, b, \dots, c$  and the error estimates  $\sigma_a, \sigma_b, \dots, \sigma_c$  of these means. Our best estimate of  $f$  is surely just  $f(a, b, \dots, c)$  using the mean values.

To estimate the error on  $f$  we note that the change in  $f$  due to small changes in  $a, b, \dots, c$  is given by

$$\Delta f = \frac{\partial f}{\partial a} \Delta a + \frac{\partial f}{\partial b} \Delta b + \dots + \frac{\partial f}{\partial c} \Delta c. \quad (12)$$

If we just averaged this expression we would get zero, since the ‘errors’  $\Delta a$ , ...  $\Delta c$  are sometimes positive, sometimes negative, and average to zero. Rather, we square the expression for  $\Delta f$ , and then average:

$$\Delta f^2 = \left(\frac{\partial f}{\partial a}\right)^2 \Delta a^2 + \dots + \left(\frac{\partial f}{\partial c}\right)^2 \Delta c^2 + \dots + 2\frac{\partial f}{\partial a}\frac{\partial f}{\partial c}\Delta a\Delta c + \dots \quad (13)$$

On average the terms with factors like  $\Delta a\Delta c$  average to zero (under the important assumption that parameters  $a$ ,  $b$ , ...  $c$  are independent). We identify the average of the squares of the changes relative to the mean values as the squares of the errors:  $\langle \Delta a^2 \rangle = \sigma_a^2$ , *etc.* This leads to the prescription

$$\sigma_f^2 = \left(\frac{\partial f}{\partial a}\right)^2 \sigma_a^2 + \dots + \left(\frac{\partial f}{\partial c}\right)^2 \sigma_c^2 + \dots \quad (14)$$

Some useful examples are

$$f = a \pm b \pm \dots \pm c \quad \Rightarrow \quad \sigma_f = \sqrt{\sigma_a^2 + \sigma_b^2 + \dots + \sigma_c^2}, \quad (15)$$

and

$$f = a^l b^m \dots c^n \quad \Rightarrow \quad \frac{\sigma_f}{f} = \sqrt{l^2 \left(\frac{\sigma_a}{a}\right)^2 + m^2 \left(\frac{\sigma_b}{b}\right)^2 + \dots + n^2 \left(\frac{\sigma_c}{c}\right)^2}, \quad (16)$$

where  $l$ ,  $m$  and  $n$  are constants that may be negative.