

# Error Analysis in Experimental Physical Science

## "Mini-Version"

by David Harrison and Jason Harlow

Last updated July 13, 2012 by Jason Harlow.

Original version written by David M. Harrison, Department of Physics, University of Toronto in January 2001. The regular version of this document is available at: <http://www.upscale.utoronto.ca/PVB/Harrison/ErrorAnalysis/index.html>

### 1. Introduction

Almost every time you make a measurement, the result will not be an exact number, but it will be a *range* of possible values. The range of values associated with a measurement is described by the uncertainty, or **error**. Since all of science depends on measurements, it is important to understand errors and get used to using them.

[An exception might be counting apples in a bowl. "Three" is an exact number of apples that does not have an error. However, if you estimate the number of apples in the back of a large apple truck and find there are 1600, this number does have an error!]

Note that in this course the word "error" does not mean "mistake"! It is better thought of as an uncertainty, meaning it quantifies how certain (or uncertain) you are about the value you measured. Basically, the error is the number that appears to the right of the  $\pm$  symbol.

We should emphasize right now that a "correct" experiment is one that has been correctly performed. *You do **not** determine the error in an experimentally measured quantity by comparing it to some number found in a book or web page!* You determine each error as part of your experiment; the value and the error are two numbers that are measured together, and you report them both, independently of what anyone else may have reported in the past.

Also, although we will be exploring mathematical and statistical procedures that are used to determine the error in an experimentally measured quantity, as you will see these are often just "rules of thumb" and sometimes a good experimentalist uses his or her intuition and common sense to simply guess what the error is. That is okay sometimes!

#### Example 1: Drawing a Histogram of repeated measurements

Imagine you have a cart on a track with a fan attached to it which causes it to accelerate along the track. You release the cart from rest and then use a digital stopwatch to measure the time it takes the cart to travel 1.50 m. You estimate the error in the distance you measured to be 1 cm, so actually the distance is  $1.50 \pm 0.01$  m. You then repeat the time measurements for a total of 30 trials. The time measurements are shown in the Table 1.



Trial #	Time (s)
1	5.55
2	5.50

3	5.56
4	5.51
5	5.49
6	5.46
7	5.49
8	5.51
9	5.32
10	5.50
11	5.49
12	5.45
13	5.30
14	5.67
15	5.64
16	5.61
17	5.66
18	5.69
19	5.38
20	5.37
21	5.74
22	5.54
23	5.56
24	5.47
25	5.49
26	5.54
27	5.45
28	5.59
29	5.52
30	5.45

**Table 1.** Time for a cart to start from rest, accelerate at a constant rate, and travel a distance of  $1.50 \pm 0.01$  m. The measurement was repeated 30 times.

- Make a histogram of the results. On the horizontal axis, plot the time from 5.0 to 6.0 seconds with tick-marks at intervals of 0.05 seconds. On the vertical axis, plot the number of trials that fall into each range.
- What is the height of the histogram? [number of trials in the most populated bin]
- What is the width of the histogram? [size of the central part of the range which contains about 2/3 of the trials.]
- Sketch a smooth bell-shaped curve through your histogram to approximate the tops of the histogram. What is the centre of this curve? [this approximates the mean].

**Example 1: Answers**

a. See the Excel spreadsheet, first sheet, available for download along with this document for work. Here is a plot of the histogram:



b. The height of the histogram is 9.

c.  $2/3$  of the 30 trials is 20. In the three bins with centres at 5.45, 5.5 and 5.55, there are a total of  $5 + 9 + 5 = 19$  trials, which is almost  $2/3$  of the total. Therefore, I would say the width of the histogram is about 3 bins, or  $5.55 - 5.45 = 0.1$  seconds.

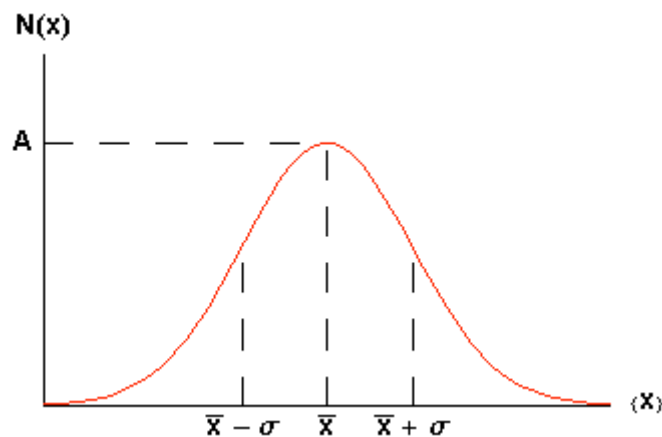
d. The smooth curve was not done in Excel, but if I was to do it by hand on a print-out, it would probably have its centre slightly to the right of 5.5 s, because the tail of this histogram extends further toward the right than the left. So perhaps the mean is about 5.51 s.

## 2. Normal Distribution

Each time you make a measurement, there is some probability that you will get a certain answer. A probability distribution is a curve which describes what the probability is for various measurements. The most important and widely used probability distribution is called the *Normal Distribution*. It was first popularized by the German mathematician Carl Friedrich Gauss in the early 1800s. It is also sometimes called the Gaussian distribution, or the bell-curve. The formula is:

$$N(x) = Ae^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

which looks like:



The symbol  $A$  is called the *maximum amplitude*.

The symbol  $\bar{x}$  is called the *mean* or *average*.

The symbol  $\sigma$  is called the *standard deviation* of the distribution. Statisticians often call the square of the standard deviation,  $\sigma^2$ , the *variance*; we will not use that name. Note that  $\sigma$  is a measure of the width of the curve: a larger  $\sigma$  means a wider curve. ( $\sigma$  is the lower case Greek letter sigma.)

Note that 68% of the area under the curve of a Gaussian lies between the mean minus the standard deviation and the mean plus the standard deviation. Similarly, 95% of the curve is between the mean minus twice the standard deviation and the mean plus twice the standard deviation.

### 3. Using the Gaussian

Suppose you make  $N$  measurements of a quantity  $x$ , and you expect these measurements to be normally distributed. Each measurement, or trial, you label with a number  $i$ , where  $i = 1, 2, 3$ , etc. You do not know what the true mean of the distribution is, and you cannot know this. However, you can estimate the mean by adding up all the individual measurements and dividing by  $N$ :

$$\text{Estimate of Mean: } \bar{x}_{\text{est}} = \frac{1}{N} \sum_{i=1}^N x_i$$

Similarly, it is impossible to know the true standard deviation of the distribution. However, we can estimate the standard deviation using our  $N$  measurements. The best estimate of the standard deviation is:

$$\text{Estimate of Standard Deviation: } \sigma_{\text{est}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_{\text{est}})^2}$$

The quantity  $N - 1$  is called the *number of degrees of freedom*. In the case of the standard deviation estimate, it is the number of measurements minus one because you used one number from a previous calculation (mean) in order to find the standard deviation.

**Example 2:** Consider the 30 measurements of time in the Table 1.

- What is the estimated mean?
- What is the estimated standard deviation?
- How many of these 30 measurements fall in the range  $\bar{x}_{\text{est}} \pm \sigma_{\text{est}}$ ? What percentage is that?

#### Example 2: Answers

- See the Excel spreadsheet, second sheet, available for download along with this document for work. The estimated mean is found to be 5.517 s.
- The estimated standard deviation is found to be 0.103 s.
- 21 of the 30 measurements fall within the range  $5.517 \pm 0.103$  s, or  $5.414 < t < 5.620$  s. This is 70% of the trials.

There is roughly a 68% chance that any measurement of a sample taken at random will be within one standard deviation of the mean. Usually the mean is what we wish to know and each individual measurement almost certainly differs from the true value of the mean by some error. But there is a 68% chance that any single measurement lies within one standard deviation of this true value of the mean. Thus it is reasonable to say that:

The standard deviation is the error in each individual measurement of the sample.

The error in a quantity is usually indicated by a  $\Delta$  (“delta”), so the above statement may be written as  $\Delta x_i = \sigma$ . This error is often called *statistical*. We shall see another type of error in the next section.

## 4. Reading Error

In the previous section we saw that when we repeat measurements of some quantity in which random statistical factors lead to a spread in the values from one trial to the next, it is reasonable to set the error in each individual measurement equal to the standard deviation of the sample. Here we discuss another error that arises when we do a direct measurement of some quantity: the reading error.

For example, imagine you use a metric ruler to measure the length of a pencil. You line up the tip of the eraser with 0, and the image below shows what you see over near 8 cm. The pencil appears to be about 8.25 cm long. What is the reading error?

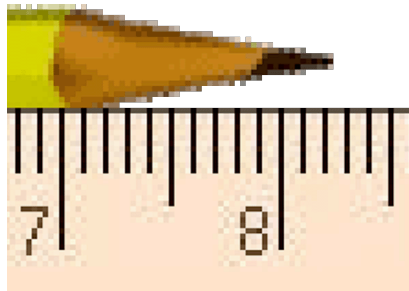


Image source: [http://www.amazingrust.com/experiments/background\\_knowledge/Measurement.html](http://www.amazingrust.com/experiments/background_knowledge/Measurement.html)

To determine the reading error in this measurement we have to answer the question: what is the minimum and maximum values that the position could have for which we will not see any difference? There is no fixed rule that will allow us to answer this question. Instead we must use our *intuition* and *common sense*!

Could the pencil actually be as long as 8.3 cm? No, I don't think so. It's clearly shorter than that. Could it be 8.28 cm? Maybe. And it could be as short as 8.23 cm, but, in my opinion, no shorter. So the range is about 8.23 to 8.28 cm. A reasonable estimate of the reading error of this measurement is half the range:  $\pm 0.025$  cm, which, to be cautious, we might round up to 0.03 cm. Then we would state that the length of the pencil is  $8.25 \pm 0.03$  cm.

For your eyes you may wish to instead associate a reading error of 0.02 cm with the position; this is also a reasonable number. A reading error of 0.04 cm, though, is probably too pessimistic. And a reading error much less than 0.01 cm is probably too optimistic.

We assume that the reading error indicates a spread in repeated measurements, just like the standard deviation discussed in the previous section. So if we get a collection of objective observers together to look at the pencil above, we expect most (ie more than 68%) of all observers will report a value between 8.22 and 8.28 cm.

Note that there is often a trade-off when assigning a reading error such as above. On the one hand we want the error to be as small as possible, indicating a precise measurement. However we also want to insure that measured value probably lies within errors of the "true" value of the quantity, which means we don't want the error to be too small.

For a measurement with an instrument with a digital readout, the reading error is " $\pm$  one-half of the last digit."

We illustrate with a digital thermometer shown to the right. The phrase " $\pm$  one-half of the last digit" above is the language commonly used in manufacturer's specification sheets of their instruments. It can be slightly misleading. It does not mean one half of the value of the last digit, which for this thermometer photo is  $0.4^{\circ}\text{C}$ . It means one-half of the power of ten represented in the last digit. Here, the last digit represents values of a tenth of a degree, so the reading error is  $1/2 \times 0.1 = 0.05^{\circ}\text{C}$ . You should write the temperature as  $12.80 \pm 0.05^{\circ}\text{C}$ .



It may be confusing to notice that we have two different specifications for the error in a directly measured quantity: the **standard deviation** and the **reading error**. Both are indicators of a spread in the values of repeated measurements. They both are describing the precision of the measurement. So you may ask: *What is the error in the quantity?*

The answer is both. But fortunately it is almost always the case that one of the two is much larger than the other, and in this case we choose the larger to be the error.

For example, if every time you measure something you always get the same numerical answer, this indicates that the reading error is dominant.

However, if every time you measure something you get different answers which differ more than the reading error you might estimate, then the standard deviation is dominant. In these cases you can usually ignore the reading error. For example, consider the rolling fan-cart we looked at in Table 1. The reading error of the digital stopwatch is  $\pm 0.005$  seconds, which is negligible compared to the standard deviation, which we found to be 0.1 seconds. Thus here the error in each measurement is the standard deviation.

Often just thinking for a moment in advance about a measurement will tell you whether repeating it will be likely to show a spread of values not accounted for by the reading error. This in turn will tell you whether you need to bother repeating it at all. If you don't know in advance whether or not you need to repeat a measurement you can usually find out by repeating it three or four times.

## 5. Significant Figures

Most introductory Physics textbooks (including the one chosen for this course) include a little section on significant figures. It provides some rules for how many significant figures to include if your number does not include an error. However, in experimental science *every number you measure or compute should include an error!* Therefore, the rules in your textbook are useless for any actual numbers measured in the Practicals. In this section, we will teach you the rules for significant figures when there is an error associated with every value.

Let's consider again the 30 data points in Table 1. The estimated standard deviation is numerically equal to 0.102933971 seconds, which is larger than the reading error for these measurements. (By "numerically" we mean that is what the calculator read when we computed the standard deviation, or what Excel produced when we used the STDEV function.) Since the estimated standard deviation is larger than the reading error, it will be the error in the value of each of the data points.

Consider one of these data points, say, the 5<sup>th</sup> trial, for which we measured 5.49 seconds. If you take the estimated standard deviation to be 0.102933971, then the data point has a value of  $5.49 \pm 0.102933971$ . What this means is that there is about a 68% chance that the true value is somewhere between 5.387066029 and 5.592933971 seconds.

Keep in mind that all of these are estimates, as we have only 30 data points, not infinity. Also, 68% is a somewhat arbitrary probability, which we have chosen to represent "most of the time". Clearly we are using WAY too many significant figures here! It would be just as instructive to say that there is about a 68% chance that the true value is somewhere between 5.4 and 5.6 seconds. Or, you could say the measurement is  $5.5 \pm 0.1$  s. In fact it is not only more concise to report this, but it is more honest.

This example illustrates two general rules for significant figures used in experimental sciences:

1. **Errors should be specified to one or two significant figures.**
2. **The most precise column in the number for the error should also be the most precise column in the number for the value.**

So if the error is specified to the 1/100th column, the quantity itself should also be specified to the 1/100th column.

**Example 3.** Express the following quantities to the correct number of significant figures:

- a.  $25.052 \pm 1.502$
- b.  $92 \pm 3.14159$
- c.  $0.0530854 \pm 0.012194$

**Example 3: Answers**

- a.  $25.1 \pm 1.5$  (or  $25 \pm 2$  is also acceptable)
- b.  $92 \pm 3$
- c.  $0.053 \pm 0.012$  (or  $0.05 \pm 0.01$  is also acceptable)



## 6. Propagation of Errors of Precision

When you have two or more quantities with known errors you may sometimes want to combine them to compute a derived number. If your errors come from standard deviation and are not reading errors, it is usually best to compute the derived number several times and use compute its standard deviation directly. However, if some reading errors are larger than the standard deviation, you can use the rules of Error Propagation to infer the error in the derived quantity.

We assume that the two directly measured quantities are  $x$  and  $y$ , with errors  $\Delta x$  and  $\Delta y$  respectively. The measurements  $x$  and  $y$  must be independent of each other.

The fractional error is the value of the error divided by the value of the quantity:  $\Delta x / x$ . The fractional error multiplied by 100 is the percentage error. Everything in this section assumes that the error is “small” compared to the value itself, i.e. that the fractional error is much less than one.

For many situations, we can find the error in the result  $z$  using three simple rules:

**Rule # 1** (sum or difference rule):

If

$$z = x + y$$

or

$$z = x - y$$

then

$$\Delta z = \sqrt{\Delta x^2 + \Delta y^2}$$

**Rule #2** (product or division rule):

If

$$z = xy$$

or

$$z = x/y$$

then

$$\frac{\Delta z}{z} = \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2}$$

Note if  $x$  is an exact number, so that  $\Delta x = 0$ , then  $\Delta z = x \Delta y$ .

**Rule #3** (exponent rule):

If

$$z = x^n$$

then

$$\frac{\Delta z}{z} = n \frac{\Delta x}{x}$$

**Example 4.** As with Example 1, you have a cart on a track with a fan attached to it, which causes it to accelerate along the track. You release the cart from rest and then use a digital stopwatch to measure the time it takes the cart to travel  $1.50 \pm 0.01$  m. You then repeat the time measurements for a total of 30

trials. The time measurements are shown in the Table 1. You model the distance as a function of time with the kinematic equation  $d = \frac{1}{2} a t^2$ . From your measurements of  $d$  and  $t$  you wish to derive  $a$ , which is  $a = 2d/t^2$ .

- Compute individual estimates of acceleration for each time measurement, using  $d = 1.50$  exactly. Estimate the mean and standard deviation of all these values of  $a$ .
- As discussed in Section 5, the best way to report the measurement for the 5<sup>th</sup> trial is:  $t = 5.5 \pm 0.1$  s. Combine this with  $d = 1.50 \pm 0.01$  m and propagate errors to compute acceleration,  $a$ .

#### Example 4: Answers

- The mean of the acceleration values is  $0.0987 \text{ m/s}^2$ , and the standard deviation is  $0.0037 \text{ m/s}^2$ .
- The equation is  $a = 2d/t^2$ , where 2 is an exact number (with no error). To compute the error, first use Rule#3 to find the error in  $z = t^2$ :  $\Delta z = 2z(\Delta t/t) = 2(5.5^2)(0.1/5.5) = 1.1$ . Then use Rule#2 to

$$\text{get the error in } y = d/z = d/t^2: \Delta y = y \sqrt{\left(\frac{\Delta d}{d}\right)^2 + \left(\frac{\Delta z}{z}\right)^2} = 0.0496 \sqrt{\left(\frac{0.01}{1.5}\right)^2 + \left(\frac{1.1}{30.25}\right)^2} = 0.0018.$$

Then use Rule#2 again with the exact number 2 as a multiplier, which simply increases the error by 2:  $\Delta a = 0.0037 \text{ m/s}^2$ . It should not be surprising to you that the error in one measurement of acceleration should be equal to the standard deviation we computed in part a.

## 7. The Error in the Mean

We have seen that when the data have errors of precision we may only estimate the value of the mean. We are now ready to find the error in this estimate of the mean.

Recall that to calculate the estimated mean we use:

$$\bar{x}_{\text{est}} = \frac{1}{N} \sum_{i=1}^N x_i$$

Each individual measurement  $x_i$  has the same error,  $\Delta x$ , which is usually the estimated standard deviation. To calculate the error in the sum of all the  $x_i$ , we use Error Propagation Rule #1 to find the error in the sum as  $\sqrt{\Delta x^2 + \Delta x^2 + \dots + \Delta x^2} = \sqrt{N\Delta x^2} = \sqrt{N}\Delta x$ . We then use Rule #2 to find the error in the sum divided by the exact number  $N$ , which gives:

$$\Delta \bar{x}_{\text{est}} = \frac{\Delta x}{\sqrt{N}}$$

#### Example 5.

- What is the mean and the error in the mean for the 30 time measurements in Table 1? [Note that the error in the mean is *not* the same as the standard deviation!]
- What is the mean and the error in the mean for the acceleration of the cart?

#### Example 5: Answers

- $5.517 \pm 0.019$  s.
- $9.867 \pm 0.067 \text{ cm/s}^2$ .