Contents

Introduction

This document is an introduction and discussion of most of the error analysis that you will need for this laboratory course. Throughout this course, you will need this packet as a reference.

Experimental Error

Essentially all experimental quantities have an uncertainty associated with them. The only exceptions are a few defined quantities such as the speed of light in a vacuum, which is defined to be 299 792 458 m/s, and the second, which is defined to be the time taken by 9 192 631 770 vibrations of the light of a specified wavelength emitted by a cesium-133 atom. In other words, we have used the fact that the speed of light and this particular cesium transition can be precisely measured, to define the meter. The uncertainty in physical measurements becomes of crucial importance when comparing experimental results with theory. The basis of the scientific method is to test our hypotheses against experimental data. It is therefore very important that you learn how to estimate how large an uncertainty or error to attach to your results due to the uncertainties in your measurements.

Note that estimating uncertainties in measured quantities is entirely different from calculating by what fraction or percentage the measured value differs from the accepted value. Even the sloppy experimenter can sometimes manage to obtain an experimental result very close to the actual value just by blind luck. Furthermore, it may be possible that either no previous measurement of the quantity exists, or the whole point of an experiment may be to test a theory which predicts a value for that quantity. In either case, it is important for the experimenter to be able to answer the following question: When someone with better equipment measures the same quantity, is the new result likely to fall within the quoted uncertainty of the present measurement?

In quoting an uncertainty, there is a standard format which is normally followed. For example, when quoting a measurement for the acceleration due to gravity g you might write:

$$g \pm \delta g = (9.83 \pm 0.05)\ \text{m/s}^2.$$

Here, g represents the best measurement you have for the acceleration due to gravity, and δg is your uncertainty in that value. This expression means that the acceleration due to gravity (as determined by the

experiment) most likely falls in the range between 9.78 m/s^2 and 9.88 m/s^2. (Later we will say more about what "most likely" means.)

The assignment of probable errors to physical data is not easy. Some sources of error can be estimated fairly accurately; other sources may be difficult or impossible to estimate. The history of physics has many notorious examples of experimenters who have grossly underestimated the errors in their measurements. This is partly the result of human nature hoping for a precise result and partly the result of lack of knowledge. Sometimes there are sources of errors that the experiments didn't know about; sometimes they knew about a source of error but didn't know how to properly estimate the effect on the experiment.

Even if you don't go into scientific research as a career, what you learn about probable errors in this course will be useful to you. We are constantly deluged with data and statistics. Sometimes it is interpreted correctly by the experts; sometimes the experts present an incomplete picture that tends to support their own opinions. In this course you will learn enough to help you understand how to evaluate data.

In the following sections, we give a brief summary of definitions and practical information on errors and treatment of data. It is a bit simplified, but it will serve as a useful introduction to the subject. If you are interested in more details, or in the derivations of the equations and conclusions presented, a list of references is given at the end of this document.

Random error. This type of error presumably arises from random effects, which are unpredictable variations in the conditions of an experiment. Random effects give rise to variations in repeated observations of the quantity being measured. Although it is not possible to compensate for the random error of a measurement result, it can usually be reduced by increasing the number of observations.

Systematic error. This type of error is often arises from systematic effects, which are usually associated with a particular instrument or experimental technique. Systematic effects occurs in every measurement, and are the same for each observation; thus systematic errors cannot be eliminated by increasing the number of observations. However, if a systematic error arises from a recognized systematic effect and can be quantified, a correction can be applied to compensate for the effect.

In most experiments, a combination of random and systematic errors are present at the same time. The distinction between random errors and systematic errors is often expressed using two words that are usually confused, accuracy and precision.

Accuracy. This is a qualitative description of how close the result of a measurement is to the true value of the quantity being measured. The accuracy of a measurement may not be necessarily be known at the time the experiment is performed. Accuracy is more strongly associated with systematic effects, although it can be affected by random effects.

Precision. This is a qualitative description of how well repeated measurements of the same quantity agree with each other. Precision is strongly associated with the smallest units that an instrument can measure as well as random effects, although it can be affected by systematic effects.

Accuracy and precision are qualitative descriptions, not quantitative descriptions. However, you can use numerical values to decide how to describe the accuracy or precision of a measurement. The generic relation between random errors, systematic errors, accuracy, and precision can be depicted in Figure 1.

For example, two students set out separately to measure the acceleration due to gravity near the Earth's surface. The standard acceleration due to gravity has a value defined as 9.806 65 m/s^2. The true value of the acceleration due to gravity may vary by as much as 0.7% of this standard and is location dependent.

$$\text{Student A:} \quad g \pm \delta g = (9.81 \pm 3.00)\ \text{m/s}^2$$
$$\text{Student B:} \quad g \pm \delta g = (8.44 \pm 0.01)\ \text{m/s}^2$$

Student A has a more accurate result than student B. However, student B has a more precise result than student A. If each student has done many trials obtain their result, then the results tell us something very important. The results of student A's trials differ widely but average to very near the true value. This may indicate that the measurement error is mainly statistical (random). The results of student B's trials are very consistent, but different from the true value. This may indicate that the measurement error is mainly systematic.

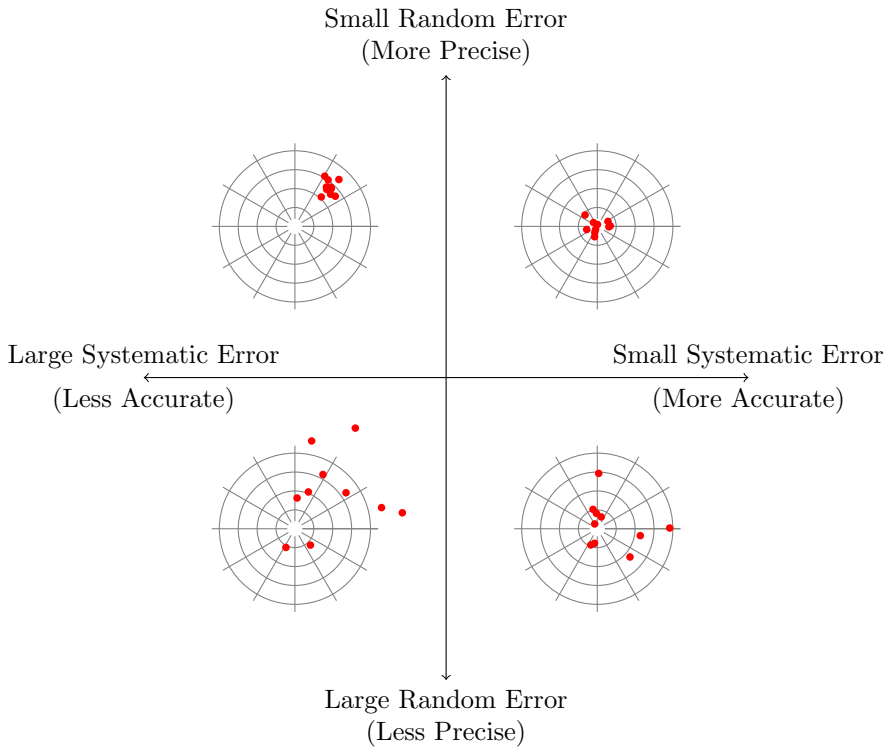


Figure 1: Examples of how systematic and random errors affect measurements and their relationship to precision and accuracy. In each quadrant of the diagram, the center of the circular target represents the the true position of an object being measured. Each dot represents the position determined by measurements of the object's position.

Uncertainties for Continuous Measurement Data — The Normal Distribution

Suppose we consider a measurement whose result can take on a continuous range of values. To be concrete, imagine a very simple experiment. You want to measure the time it takes a ball to fall 1.020 m. To get an accurate value you use a computerized timer that reports values to three decimal places and repeat the measurement 200 times. Figure 2 shows the time data collected from our hypothetical experiment in the form of a histogram. The vertical height of each bar gives the number of measurements that fell within the range of the bin. For example, there were 13 measurements with a falling time between 0.435 s and 0.440 s. The distribution in Figure 2 is typical of what real data from a well designed and executed experiment might look like. The most notable features are:

1. Measurements with values near the center of the distribution occur more often than measurements with values near the edges of the distribution.

2. The distribution is reasonably symmetric about its center. There is no obvious skewing toward the high or low side.

If we continued to make measurements and made the bins very fine, our histogram might begin to look like the smooth bell-shaped curve. This curve is the limiting case in an ideal situation. It is referred to as the normal distribution or Gaussian distribution. Measurement errors that follow this distribution are said to be normally distributed. The mathematical theory of the Gaussian curve is important, but much of it is beyond the scope of this handout. What is important is that if the fluctuations in a given quantity

are random, then a distribution resembling a normal distribution is usually found. Then the results of the mathematical theory can be used to estimate uncertainties in our data.

The normal distribution is can be described by its center and its spread. The probability density of a normal distribution has a characteristic function of a general form.

$$p(x|\bar{x}, s_x) = \frac{1}{s_x\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2s_x^2}\right) \tag{1}$$

The statistic $\bar{x}$ (x-bar) is the center of the distribution, and is called the mean. The statistic s_x is the spread of the distribution, and is called the standard deviation.

If you wish to overlay the normal distribution over a histogram of the related data, then you need a factor of bn to the match the scale of the histogram.

$$f(x|\bar{x}, s_x, n, b) = bnp(x|\bar{x}, s_x) \tag{2}$$

The statistic n is the number of observations in the sample. The number b is the width of the bins used in your histogram.

Mean (of the sample). This is one way of describing the center of a distribution, another is the median. If you have a sample with n observations of a quantity x, the mean $\bar{x}$ is found by taking the (arithmetic) average of all the values in the sample $(x_1, x_2, x_3, \ldots, x_n)$.

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + x_3 + \cdots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

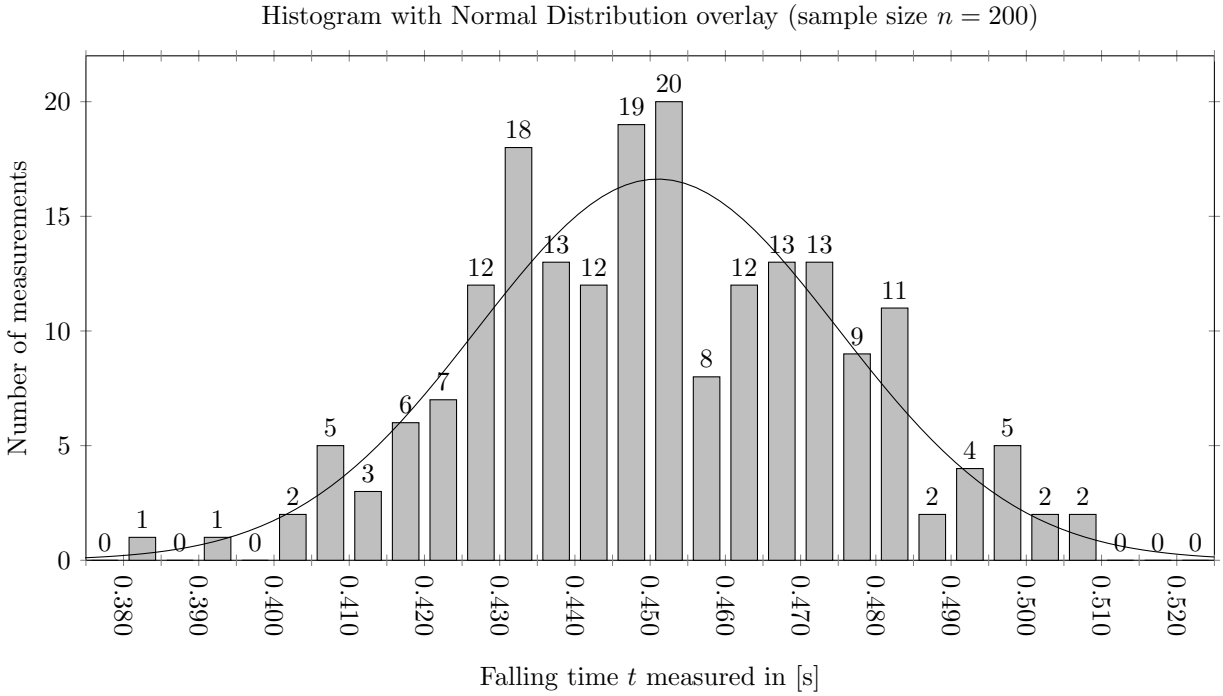


Figure 2: This graph uses 200 points of data that has been grouped into bins of width 0.005 s. The mean value of the data is 0.451 s, and the standard deviation is 0.024 s. The normal distribution curve using these statistics has been overlayed on the histogram.

Standard deviation (of the sample). This is one way of describing the spread of a distribution. The calculation for the standard deviation of a sample requires you know the mean. First, for each observation in the sample find the deviation from the mean (you will now have a list of deviations). Second, square each deviation (you will now have a list of squares). Third, add the squares (you will now have a a single value). Fourth, divide the sum by one less than the original sample size. Fifth, take the square root of the quotient.

$$s_x = \sqrt{\left(\frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)} \tag{4}$$

The standard deviation provides a way to express the standard uncertainty of a measurement.

- $\approx 68\%$ of observations fall within 1 standard deviation of the mean ($\bar{x} \pm s_x$).
- $\approx 95\%$ of observations fall within 2 standard deviations of the mean ($\bar{x} \pm 2s_x$).
- $\approx 99.7\%$ of observations fall within 3 standard deviations of the mean ($\bar{x} \pm 3s_x$).

Often, you won't be interested how an individual observation will vary from the mean. Rather, you will need to know how the mean itself can vary with repeated iterations of the experiment. Each time an experiment is performed, the sample data will consist of different values. The mean of each iteration of the experiment will likely also be different. The set of all means (from normally distributed data) are themselves normally distributed. Since it is impractical to run endless iterations of an experiment, the spread in the mean values can be estimated.

If you could collect all the measurement data for a particular quantity in a particular experimental setup, then you would have the entire population of data. However, in experiments such as those in this course it is always possible (at least in principle) to get another measurement for a particular experimental setup. Thus your sample can always be made larger, but never the reach the size of the entire population of possible measurements.

Mean (of the population). The parameter μ_x is the mean of the entire population of possible measurements. Because you cannot make endless measurements, the population mean μ_x is generally not known.

Standard deviation (of the population). The parameter σ_x is the standard deviation of the entire population of possible measurements. Because you cannot make endless measurements, the population standard deviation σ_x is generally not known.

Standard uncertainty in the mean. The statistic u_x is the standard uncertainty in the mean. It describes the spread of mean values taken from multiple iterations of an experiment. This is also known as the standard deviation of the mean or the standard error of the mean. Be careful not to confuse it with the standard deviation of the sample. The standard uncertainty in the mean u_x is related to the standard deviation of the sample s_x and the number of observations n.

$$u_x = \frac{s_x}{\sqrt{n}} \tag{5}$$

There is an alternate notation taken from the calculus of variations that will be frequently used in this course.

$$\delta x = u_x \tag{6}$$

The standard uncertainty in the mean can be used to construct a confidence interval similar to those that use the standard deviation of the sample. In PHSX 216N and PHSX 218N, an expression for the range of most likely values is all that will be required. A more rigorous statistical treatment of confidence intervals does exist for the treatment of the standard uncertainties, even including confidence levels (expressed as percentages).

$$\bar{x} \pm \delta x \tag{7}$$

The above discussion assumes that each of the measurements have equal intrinsic accuracy and are therefore equally weighted in the determination of the mean. If some of the measurements are more or less certain than others, more advanced techniques exist for including a weighting factor when determining each statistic. These techniques are discussed in the references.

Example — Falling Time

Since the measured time uses the variable t instead of the variable x, an appropriate change in notation should be made.

$$(x, \bar{x}, s_x, u_x) \mapsto (t, \bar{t}, s_t, \delta t) \tag{8}$$

The data used to construct the histogram in Figure 2 has a mean time $\bar{t} = 0.451$ s with a standard deviation $s_t = 0.024$ s. The sample has size $n = 200$. The uncertainty in the mean can then be calculated.

$$\delta t = u_t = \frac{s_t}{\sqrt{n}} = \frac{(0.024 \text{ s})}{\sqrt{(200)}} = 0.0016970\ldots \text{ s} \approx 0.002 \text{ s} \tag{9}$$

One you know the standard uncertainty in the mean, you can a range of most likely values.

$$\bar{t} \pm \delta t = (0.451 \pm 0.002) \text{ s} = \begin{cases} 0.453 \text{ s} \\ 0.449 \text{ s} \end{cases} \tag{10}$$

The falling time for a ball dropped from a height of 1.020 m is most likely between 0.449 s and 0.453 s.

Uncertainties for Integer Measurement Data — The Square Root Rule (Poisson Distribution)

Often the result of an experiment or measurement is an integer. For example, the number of mice out of an initial sample that die within one year or the number of radioactive nuclei out of a sample that decay in one second. The standard deviation of the number of such "events" (i.e., deaths, decays, etc.) can be estimated by the square root rule.

Suppose you have a sample of size n, events occur at a rate of λ, and you make measurements over a time interval t. The mean number of events can be calculated from a known event rate, found directly from a single iteration of the experiment, or averaged from multiple iterations of the experiment.

$$\bar{x} = \lambda t \qquad \text{or} \qquad \bar{x} = \begin{bmatrix} \text{number of events} \\ \text{from 1 experiment} \end{bmatrix} \qquad \text{or} \qquad \bar{x} = \begin{bmatrix} \text{average number of events} \\ \text{from repeated experiments} \end{bmatrix} \tag{11}$$

The standard deviation of the sample and the standard uncertainty in the mean are equal. Be aware that they are defined differently than for a Gaussian distribution.

$$u_x = s_x = \sqrt{\bar{x}} \tag{12}$$

For this to be an accurate estimate, the following conditions must be satisfied. (The better they are satisfied the better the estimate of u_x.)

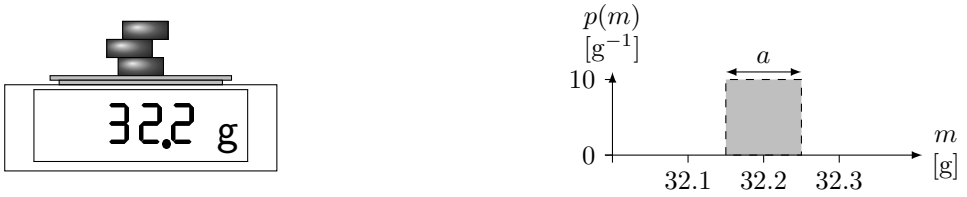
1. The number of events $\bar{x}$ must be large. (Some would argue that $\bar{x} > 10$ is enough.)
2. The probability that one specific event (i.e. the death of a specific mouse, or the decay of a specific nuclei) occurring during the measurement must be small.
   - If we did an experiment to see how many mice out of 100 will die in 75 years, the answer would be that all the mice would surely die (100 mice ± 0 mice). The uncertainty would not follow the square root rule because the chances of any specific mouse dying are not small.
   - If we start with 10^8 radioactive nuclei which decay at a rate of 10^3 nuclei per second, in a 10 second experiment you can expect $\bar{x} \pm u_x = (10\,000 \pm 100)$ nuclei to decay. The square root rule should work very well because $\bar{x} \gg 1$ and the probability of any given nucleus decaying is 10^{-4} during the experiment.

Uncertainties for a Single Measurement — Digital Scale

Often there will be quantities you will only measure once for an experiment, such as the length of a string or mass of a cylinder. In these cases, the statistics of the Gaussian distribution, the Poisson distribution, or any other distribution are inappropriate because a single measurement does not describe a distribution.

When you make a measurement using a digital instrument, the display has a smallest place value. It cannot show any number smaller. Any measured quantity that has a fractional part smaller than what the digital scale can display will be rounded.

Figure 3 shows an example with an electronic scale measuring the mass of some objects as 32.2 g. The probability density associated with the measurement is rectangular because without more information the actual mass of the objects could be any value between 32.15 g and 32.25 g with equal probability. Any value within this range is rounded to 32.2 g.



(a) Digital scale reading of the mass of some objects.

(b) Probability density p of the scale reading x.

Figure 3: The reading from a digital scale and the probability density of width a associated with the measurement of the displayed reading.

Standard Uncertainty (of a digital measurement) In general, the probability density for any digital reading will have a center value which is the displayed reading. The probability density will have a width a where all values are equally likely and are rounded to the center value. The width a will often be equal to the size of the smallest measurable unit of the display. The standard uncertainty of a measured value x with a rectangular probability distribution is

$$u_x = \frac{a}{2\sqrt{3}}. \tag{13}$$

The confidence interval $x \pm u_x$ has a confidence level of 58%.

The probability distribution in Figure 3 has the width $a = 0.1$ g. The measured quantity uses the variable m instead of x, so the standard uncertainty variable will be chosen as $\delta m = u_m$. The standard uncertainty and a range of most likely values can then be calculated.

$$\delta m = u_m = \frac{a}{2\sqrt{3}} = \frac{(0.1\text{ g})}{2\sqrt{3}} \approx 0.028867\text{ g} \approx 0.03\text{ g} \tag{14}$$

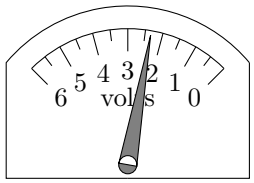
$$m \pm \delta m = (32.2\text{ g}) \pm (0.03\text{ g}) = \begin{cases} 32.23\text{ g} \\ 32.17\text{ g} \end{cases} \tag{15}$$

The total mass of the three objects is most likely between 32.17 g and 32.23 g.

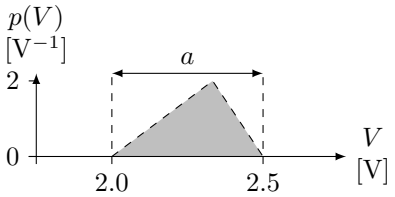
Uncertainties for a Single Measurement — Analog Scale

Another type of instrument you might use for a single measurement would be one with an analog scale. These include needle gauges and meter sticks. Generally, a range of measurable values is visible and you have to pick a physical position on the device to read the measurement.

Where you would read the measurement is the most probable value. The probability of it being different drops as you move along the scale to either side, and it will eventually reach zero. A fair estimate of where



(a) Analog scale reading of an electric potential.



(b) Probability density p of the voltmeter reading V.

Figure 4: The reading from a digital scale and the probability density of width a associated with the measurement of the displayed reading.

the probability reaches zero is at the marks on the scale on either side of your measurement. In Figure 4, you can see this creates a triangle probability distribution.

Standard Uncertainty (of an analog measurement) A triangular probability density will have a base of width a. The value of a may often be equal to the smallest increment of the scale. The standard uncertainty of a measured value x with a triangular probability distribution is

$$u_x = \frac{a}{2\sqrt{6}}. \tag{16}$$

The confidence interval $x \pm u_x$ has a confidence level of 65%.

The probability distribution in Figure 4 has a base of width $a = 0.5$ V. The measured quantity uses the variable V instead of x, so the standard uncertainty variable with be chosen as $\delta V = u_V$. Unlike the digital scale, the analog scale requires your judgment to report a value. You will need to use your best estimate to judge the measurement value. In this example, the reading might be $V = 2.3$ V (although you might estimate it to be 2.4 V). The standard uncertainty and a range of most likely values can then be calculated.

$$\delta V = u_V = \frac{a}{2\sqrt{6}} = \frac{(0.5\text{ V})}{2\sqrt{6}} \approx 0.102062\text{ V} \approx 0.1\text{ V} \tag{17}$$

$$V \pm \delta V = (2.3\text{ V}) \pm (0.1\text{ V}) = \begin{cases} 2.4\text{ V} \\ 2.2\text{ V} \end{cases} \tag{18}$$

The electric potential is most likely between 2.2 V and 2.4 V.

Error Propagation — Uncertainties of Calculated Values

Usually you cannot make a direct measurement of the quantity you are interested in. You must measure other quantities and then calculate the desired quantity from them. In the example of the falling ball (Figure 2), you might be interested in the acceleration due to gravity g. It can be calculated from the falling time t and the falling distance h.

$$g = \frac{2h}{t^2} \tag{19}$$

Determining the uncertainty in g can be done if the uncertainties in t and h are both known.

The rules for error propagation can be readily derived using calculus. Some results here are stated here. The quantity to be calculated is Q and its uncertainty is δQ. The value of Q will depend on measured quantities A and B, which have uncertainties δA and δB.

Rule #1. If $Q = cA$ where c is a constant (or a quantity with negligible uncertainty), then

$$\frac{\delta Q}{|Q|} = \frac{\delta A}{|A|} \quad \text{or equivalently} \quad \delta Q = |c|\delta A \tag{20}$$

Rule #2. If $Q = cA^m$ where c and m are constants, then

$$\frac{\delta Q}{|Q|} = |m|\frac{\delta A}{|A|} \quad \text{or equivalently} \quad \delta Q = |cmA^{(m-1)}|\delta A \tag{21}$$

Rule #3. If $Q = A + B$ or $Q = A - B$, then

$$\delta Q = \sqrt{(\delta A)^2 + (\delta B)^2} \tag{22}$$

Rule #4. If $Q = cA^m B^n$ where c, m, and n are constants, then

$$\frac{\delta Q}{|Q|} = \sqrt{\left(m\frac{\delta A}{A}\right)^2 + \left(n\frac{\delta B}{B}\right)^2} \tag{23}$$

"Brute Force" If the functional form for Q is not covered by Rules #1, #2, #3, or #4, you can calculate the uncertainty in Q by the "brute force" method. Calculate Q using the mean value of A, then recalculate Q using $A + \delta A$. The positive difference between the two results will give δQ.

$$\text{If} \quad Q = f(A) \quad \text{then} \quad \delta Q = |f(A) - f(A + \delta A)| \tag{24}$$

If Q is a function of more than one variable, vary each separately. Then combine the separate terms in quadrature. For example, if $Q = f(A, B, \ldots)$, then the uncertainty in Q is

$$\begin{aligned} \text{If} \quad Q = f(A, B, \ldots) \quad \text{then} \quad \delta Q_A &= |f(A, B, \ldots) - f(A + \delta A, B, \ldots)| \\ \delta Q_B &= |f(A, B, \ldots) - f(A, B + \delta B, \ldots)| \\ &\vdots \\ \delta Q &= \sqrt{(\delta Q_A)^2 + (\delta Q_B)^2 + \cdots} \end{aligned} \tag{25}$$

Often one or two of the terms use to calculate the uncertainty δQ will dominate over the others; the others will give negligible contribution to δQ. Typically, if the dominate term comes from a single measurement from a digital or analog scale, then the experiment should probably be redesigned.

Error Propagation — Adding and Subtracting Measurement Data

Many times you will need to measure something and your measuring instrument will not be sufficient to measure the entire quantity in one step. This may mean you must measure it in pieces and add the pieces together to get a best value. You might need to measure the mass of some metal disks, but the total mass exceeds the limit of your scale. You will separate the disks into smaller groups then add the masses of the groups together to get the total mass of the disks.

You must use Rule #3 to find the uncertainty of the total value of all the measurements. Rule #3 is expandable to accommodate any number of measurements you may need to make. Thus the uncertainty of the total measurement will not be the same as the standard uncertainty of one measurement.

$$\text{If} \quad x = x_a + x_b + \cdots \quad \text{then} \quad \delta x = \sqrt{(\delta x_a)^2 + (\delta x_b)^2 + \cdots} \tag{26}$$

Rule #3 also applies if you need to subtract one measurement from another to get the value you want. For example, you might need to find the mass of an amount of liquid by subtracting the mass of the empty container from the filled container. Or you may need to find the distance between two positions (this is a common type of measurement and an example is in the next section).

Example — Falling Height

The data for Figure 2 was collected using a height of 1.020 m. The uncertainty for this value however, is NOT that of a single analog measurement. This is because the initial position of the ball and final position of the ball are two separate measurements. In fact, any length measurement will be the difference between two positions. Because of this, the uncertainty of a length must be determined by Rule #3 from the position uncertainties.

$$\text{Length measurement} \qquad \ell = x_2 - x_1 \tag{27}$$

$$\text{Length uncertainty (apply Rule \#3)} \qquad \delta\ell = \sqrt{(\delta x_2)^2 + (\delta x_1)^2} \tag{28}$$

Even if one of the positions is measured to be zero, or set to zero, or made zero by any means, it still has an uncertainty associated with it.

Suppose the measuring stick to determine the positions of the falling ball was only reliable read to the nearest 0.01 m. (Even though meter sticks are usually marked to the nearest 0.001 m, if you can't reliably the The standard uncertainty for both positions could then be found, and then the uncertainty for the falling height.

$$\delta y_1 = \frac{a}{2\sqrt{6}} = \frac{(0.01\text{ m})}{2\sqrt{6}} \approx 0.002041\text{ m} \tag{29}$$

$$\delta y_2 = \frac{a}{2\sqrt{6}} = \frac{(0.01\text{ m})}{2\sqrt{6}} \approx 0.002041\text{ m} \tag{30}$$

$$\delta h = \sqrt{(\delta x_2)^2 + (\delta x_1)^2} \approx \sqrt{(0.002041\text{ m})^2 + (0.002041\text{ m})^2} \approx 0.002887\text{ m} \approx 0.003\text{ m} \tag{31}$$

Using this, you can now find the interval where the falling height is most likely to be.

$$h \pm \delta h = (1.020\text{ m}) \pm (0.003\text{ m}) = \begin{cases} 1.023\text{ m} \\ 1.017\text{ m} \end{cases} \tag{32}$$

The falling height is most likely between 1.017 m and 1.023 m.

Example — Acceleration due to Gravity

The acceleration due to gravity has an expression that at first glance does not appear to match any of the Rules for error propagation. But denominators can be rewritten as numerators with a change of sign on their exponents.

$$g = \frac{2h}{t^2} = 2ht^{-2} \tag{33}$$

Now it matches the form of Q that goes with Rule #4.

$$\begin{array}{rcccc} Q & = & c & A^m & B^n \\ g & = & 2 & h^1 & t^{-2} \end{array} \tag{34}$$

The constant are $c = 2$, $m = 1$, and $n = -2$.

$$\frac{\delta g}{g} = \sqrt{\left(1\frac{\delta h}{h}\right)^2 + \left(-2\frac{\delta t}{t}\right)^2} \tag{35}$$

Without even knowing what the measurements are, this expression is the generic form for calculating the uncertainty in the acceleration due to gravity from the falling time and falling height for an object dropped from rest.

In general, you shouldn't calculate the desired quantity for each trial of an experiment and then combine them (by averaging or some other technique). Since the time data used to create Figure 2 has a mean value,

you can use that to calculate the most probable value for the acceleration due to gravity. Stated again, the relevant values are $\bar{t} = 0.451$ s and $h = 1.020$ m.

$$g = \frac{2h}{\bar{t}^2} = \frac{2(1.020\text{ m})}{(0.451\text{ s})^2} = 10.029449\ldots\ \text{m/s}^2 \tag{36}$$

$$\delta g = g\sqrt{\left(1\frac{\delta h}{h}\right)^2 + \left(-2\frac{\delta t}{\bar{t}}\right)^2} = (10.12778\text{ m/s}^2)\sqrt{\left(\frac{0.003\text{ m}}{1.030\text{ m}}\right)^2 + \left(2\frac{0.002\text{ s}}{0.451\text{ s}}\right)^2} \tag{37}$$
$$\delta g = 0.093717\ldots\ \text{m/s}^2 \approx 0.10\ \text{m/s}^2$$

Thus the result can be stated as

$$g \pm \delta g = (10.03 \pm 0.10)\ \text{m/s}^2 \tag{38}$$

The range of this interval is from 9.93 m/s^2 to 10.13 m/s^2. You may notice that the defined value for the standard acceleration due to gravity (9.806 65 m/s^2) is outside this range. This could imply the standard is poorly defined or it doesn't apply to the location the experiment was performed. It is more likely either the experiment was performed poorly or there is a systematic error that has reduced the accuracy of the experiment.

Rounding, Significant Figures, and Scientific Notation

Do not round numbers until you arrive at a final result. Rounding between steps when calculating can introduce mistakes in your calculations. If you do have to round between steps, be sure to carry several extra digits to avoid rounding errors. Be aware that some operations are more sensitive to rounding than others.

In general, most uncertainties will only have one or two significant figures. Unlike usual rounding procedures where you can round up or down, depending on the value you are rounding, uncertainties should always be rounded up.

The number of significant figures you give in any result ($x \pm \delta x$) should be consistent. This means the smallest place value in both numbers is the same. It doesn't make sense to give too many significant figures or too few significant figures.

incorrect:	9.5876543 ± 0.67789798	10 ± 0.0141015
incorrect:	9.5876543 ± 0.68	10 ± 0.015
correct:	9.59 ± 0.68	10.000 ± 0.015

Additionally, if you express your values in scientific notation, then they should have the same power of ten. This means the mean value will be in proper scientific notation, however the uncertainty will likely not be in proper scientific notation. The underlying reason is ease of readability. Do NOT use computer E notation unless you are communicating with a computer or discussing how to enter numbers into a computer.

incorrect:	$1.6 \times 10^{-19} \pm 1.2 \times 10^{-21}$	WRONG:	1.6E−19 ± 1.2E−19
incorrect:	$1.602 \times 10^{-19} \pm 1.2 \times 10^{-21}$	WRONG:	1.602E−19 ± 1.2E−19
correct:	$1.602 \times 10^{-19} \pm 0.012 \times 10^{-19}$	WRONG:	1.602E−19 ± 0.012E−19
correct:	$(1.602 \pm 0.012) \times 10^{-19}$	WRONG:	(1.602 ± 0.012) E−19

Making Graphs

Graphs are a very useful means of exhibiting data so that it can be readily visualized. Graphs can be used for smoothing out data and often provide a convenient means of analyzing data to provide a "best fit" value of some physical quantity. It is assumed you have had some experience with graphing data and this section will deal mainly with the finer points.

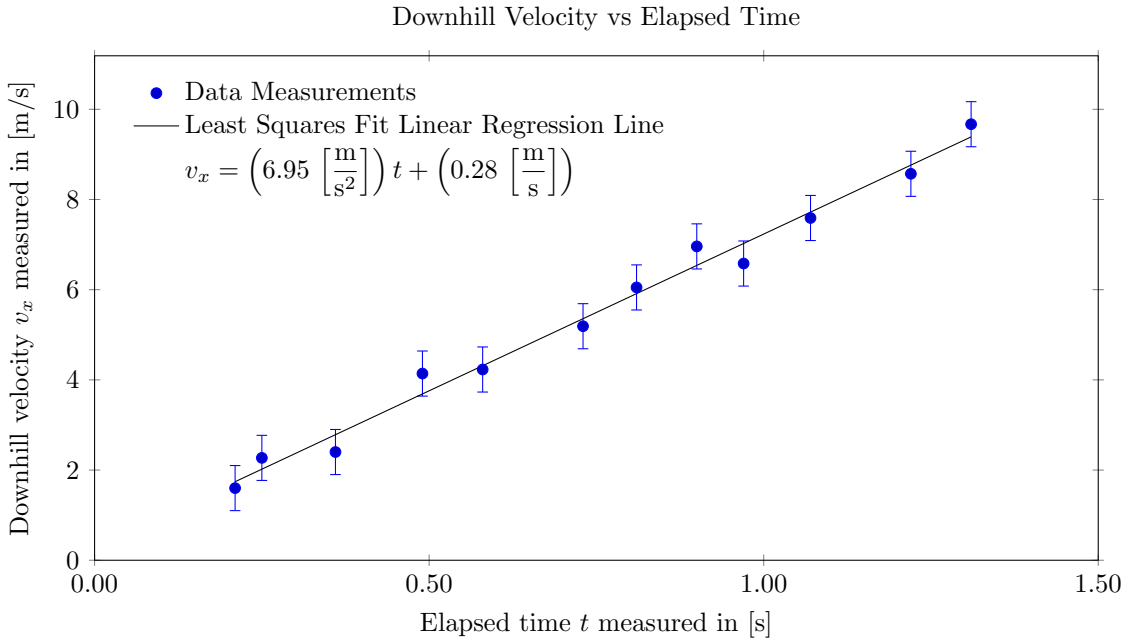


Figure 5: Example of a plot with error bars. A linear regression line is fit to the data using the method of least squares.

Choosing the Scale

The scale of your graph is often a matter of usefulness and aesthetics. Often you will need to make more than one attempt to find a scale that is satisfactory. The resolution of your graph should reflect the uncertainty of the data. If the uncertainties are on the order of 1% of the measured values, then you will want a finer resolution to the graph's scale; whereas if they are on the order of 20% then the scale can be more coarse. If you were to analyze the data based on the visual representation of the data, then the scale can greatly affect your results.

Error Bars

Usually the quantities you are plotting will have an uncertainty associated with them. The uncertainty is indicated by drawing a line through the point to indicate the uncertainty. For example, if the velocity of an object is determined to be $v_x \pm \delta v_x = (1.80 \pm 0.15)$ m/s, you would plot the point at 1.80 m/s and the vertical bar would extend from 1.65 m/s to 1.95 m/s. Figure 5 gives an example of linear experimental data with error bars. If there is uncertainty in both the horizontal and vertical coordinate, you would generally show this with both horizontal and vertical error bars. Usually the error bars indicate one standard deviation or one standard uncertainty.

The Best Fit Curve

Perhaps in middle school, high school, or other college courses you did data analysis by drawing lines by hand. However, in this course all data will be analyzed by numerical means, such as a linear regression.

In practice, graphical analysis is often used to see if your data are consistent with a theory and perhaps to determine the value of some constant in the theoretical expression which best fits your data. For example, an object dropped from rest is expected to have a speed at time t of

$$v = gt \tag{39}$$

where g is the acceleration due to gravity. This equation describes the speed as a linear function of time. The value of g could be determined by graphing different experimental values for v and t then determining the

slope of the resulting line. Suppose, however, that your data consisted of various values of falling distances h and falling times t of a ball dropped from rest. Then we expect

$$h = \frac{1}{2}gt^2 \tag{40}$$

This is the equation for a parabola which would be more difficult to fit graphically than a linear equation. In fact, many functions you might wish to model are not linear. Since solving linear equations is usually easier than non-linear equations, the are techniques for re-expressing various classes of equations as linear equations. Linearizing a function means to re-express the parameters in the equation so that the resulting function describes a straight line. Some functions cannot be linearized and must be solved with advanced techniques. However, you will find with most of the labs for PHSX 216N and PHSX 218N linear fits are possible.

Linearizing Quadratic Functions

For a quadratic function, linearizing requires no special mathematical transformations You can group the parts of the equation so they correspond to the different parts of the slope-intercept form of the equation of a line.

$$\begin{array}{lccccccc} \text{generic equation of a line} & y & = & m & x & + & b \\ \text{meaning of each variable} & \begin{bmatrix}\text{vertical}\\ \text{variable}\end{bmatrix} & = & \begin{bmatrix}\text{slope of}\\ \text{the line}\end{bmatrix} & \begin{bmatrix}\text{horizontal}\\ \text{variable}\end{bmatrix} & + & \begin{bmatrix}\text{vertical}\\ \text{intercept}\end{bmatrix} \\ & h & = & \left(\frac{g}{2}\right) & \left(t^2\right) \end{array} \tag{41}$$

Now the squared time t^2 is the quantity to be plotted on the horizontal axis. After determining the slope of line fit to the data, the acceleration due to gravity can then be found.

$$[\text{slope}] = \frac{g}{2} \tag{42}$$

$$2 \times [\text{slope}] = g \tag{43}$$

You will notice the value of the intercept is predicted to be zero before you even perform the experiment. After finding the best fit line, you will discover the intercept is unlikely to be zero. The size of the deviation between the experimental value of the intercept and the theoretical value of the intercept can sometimes be used as a clue to the accuracy of the experiment.

Linearizing Exponential Functions

In some cases the quantity that is to be determined by experiment is in the exponent. Two common examples of this are radioactive decay (exponential) and period of a simple pendulum (power law). In radioactive decay, the number of counts per second R is expected to vary with time as

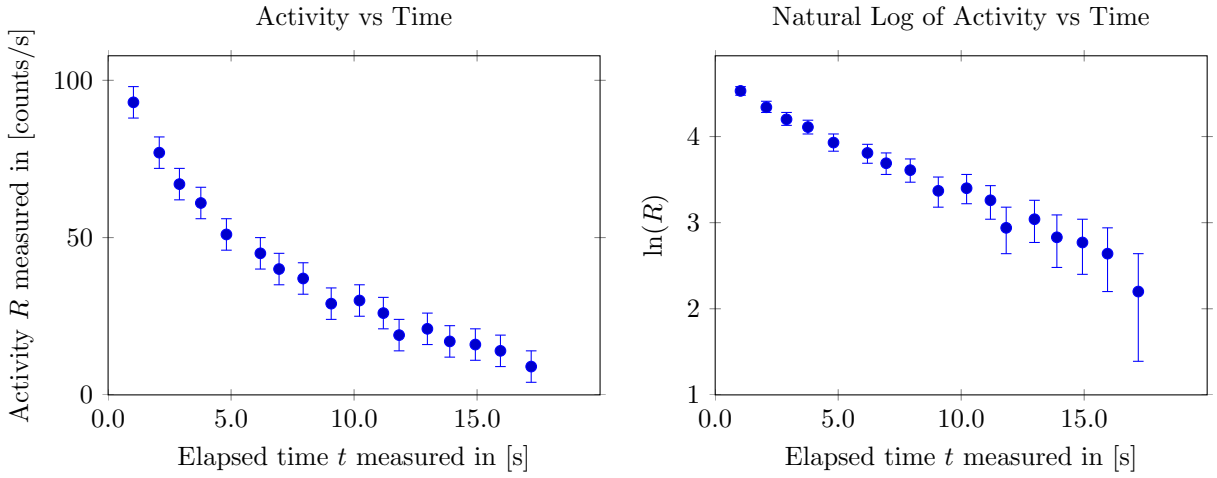
$$R = Ce^{-\lambda t} = C\exp\left(-\lambda t\right) \tag{44}$$

where C and λ are quantities to be determined by measuring R as a function of time. In Figure 6, typical data for this type of experiment are shown. Fitting the data directly with an exponential equation to determine λ would be tedious. However, taking the natural logarithm of both sides of the equation results in an expression that can be simplified to something less cumbersome.

$$\ln(R) = \ln\left(C\exp(-\lambda t)\right) \tag{45}$$

$$\ln(R) = \ln(C) + \ln\left(\exp(-\lambda t)\right) \tag{46}$$

$$\ln(R) = \ln(C) + (-\lambda t) \tag{47}$$



(a) Exponential plot from raw data.

(b) Linear Plot from transformed data.

Figure 6: A exponential plot can be transformed into a linear plot. Note that the error bars also transform.

This equation has the same form as the slope-intercept form of the equation of a line. The slope of this line in equal to $-\lambda$ and the vertical intercept is equal to $\ln(C)$.

$$\begin{array}{ccccccc} \ln(R) & = & (-\lambda) & t & + & \ln(C) \\ y & = & m & x & + & b \end{array} \tag{48}$$

Figure 6 gives an example of this transformation. The error bars must also be found for the newly transformed values. Note that the error bars in the new plot are not symmetric.

$$\text{Old upper bound: } R+\delta R \qquad \text{New upper bound: } \ln(R+\delta R) \tag{49}$$

$$\text{Old lower bound: } R-\delta R \qquad \text{New lower bound: } \ln(R-\delta R) \tag{50}$$

Linearizing Power Law Functions

A power law is a function is of the form

$$y = ax^m \tag{51}$$

where the coefficient a and the exponent m are to be determined. Just as with exponential function, taking the natural logarithm of both sides proves to be useful.

$$\ln(y) = \ln(ax^m) \tag{52}$$

$$\ln(y) = \ln(a) + \ln(x^m) \tag{53}$$

$$\ln(y) = \ln(a) + m\ln(x) \tag{54}$$

Again, the equation now has a form that corresponds to the slope-intercept form of the equation of a line. The slope of the line is equal to n and the vertical intercept is equal to $\ln(a)$.

$$\begin{array}{ccccccc} \ln(y) & = & m & \ln(x) & + & \ln(a) \\ y & = & m & x & + & b \end{array} \tag{55}$$

Again, the error bars in the new plot are not symmetric about the central value.

The Linear Fitting of Data — Method of Least Squares

Once you've decided that your data can be fit with a line, you want to get the "best" fit possible. There must be some criterion by which a "best" fit can be determined. Although there are many criteria to chose from, most line fitting algorithms you will likely encounter use the method of least squares. Some of the details are briefly explained here, however you can find more in depth material in the references.

General approach

Suppose you're trying to fit the linear function $f(x) = mx + b$ to a sample of n data points $(x_i, y_i, \delta y_i)$ where each y_i has an uncertainty δy_i and there is no uncertainty in x_i (or the uncertainty is negligible). The best choices for the coefficients m and b are those for which the discrepancy between the values of the measurements y_i and the corresponding values $f(x_i)$ is minimized.

$$[\text{discrepancy}]_i = y_i - f(x_i) \tag{56}$$

If the uncertainties of your data are not all equal, then they should have differing importance on your calculations. How much a value influences your calculations is referred to as a weight (by analogy to how differing masses have different physical weight). Values you are more certain about have a greater weight and values you are less certain about have a lesser weight. You can weigh each discrepancy by dividing it by the uncertainty of the measurement.

$$\begin{bmatrix} \text{weighted} \\ \text{discrepancy} \end{bmatrix}_i = \left(\frac{y_i - f(x_i)}{\delta y_i} \right) \tag{57}$$

Adding all the weighted discrepancies will give you the total discrepancy. However, because some are positive and some are negative there are technical issues that arise. So, it is customary to square each weighted discrepancy then add them together. Furthermore, as our data in this course will typically follow a normal distribution and the typical data point is within 1 standard deviation of the mean, each squared weighted discrepancy will be set equal to 1.

$$1 = \left(\frac{y_i - f(x_i)}{\delta y_i} \right)^2 = \left(\frac{y_i - mx_i + b}{\delta y_i} \right)^2 \tag{58}$$

Therefore the sum of the weighted values should be about equal to the number of data points n.

$$n = \sum_{i=1}^{n} \left(\frac{y_i - f(x_i)}{\delta y_i} \right)^2 = \sum_{i=1}^{n} \left(\frac{y_i - mx_i + b}{\delta y_i} \right)^2 \tag{59}$$

You may or may not have a lot of experience minimizing functions. To minimize a function, you will need to find the derivatives with respect to each of the free parameters (in this case m and b). Because the minimum point of a function has a slope of zero, the derivatives you calculated are each set equal to zero.

$$\frac{dn}{dm} = 0 \qquad \text{and} \qquad \frac{dn}{db} = 0 \tag{60}$$

Solving this system of two equations will give you values for m and b. You will not be required to do this calculation yourself. You will use the aid of a computer program to find the solutions.

Solution with equal weighted data

When all the uncertainties δy_i are equal, the resulting solutions to the slope m and intercept b have a relatively clean form. First, calculate the mean of the x-values and the mean of the y-values.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \text{and} \qquad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{61}$$

The calculation for the slope of the line is a bit more involved.

$$m = \frac{\sum_{i=1}^{n} \Big((x_i - \bar{x})(y_i - \bar{y}) \Big)}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{62}$$

The intercept of the line can be found once you know the means and the slope.

$$b = \bar{y} - m\bar{x} \tag{63}$$

As stated previously, you will usually be using a computer program to calculate these for you.

Uncertainties

To determine the uncertainty in the slope δm and the uncertainty in the intercept δb you will need to apply methods of propagation of uncertainties. This calculation is beyond the scope of this course; you will be able to determine these uncertainties with the aid of a computer program. It will give you the fitted slope $m \pm \delta m$ and fitted intercept of the line $b \pm \delta b$.

The uncertainties in the slope and intercept depend heavily on the uncertainties in the data, so it is important that the uncertainties in the data are good estimates of the actual uncertainties in the measurement. One important assumption that is made in this analysis is that the uncertainties in the independent variable x are negligible compared to the uncertainties in the dependent variable y.

- If the errors in the independent variable are comparable to those in the dependent variable, then another analysis method is required.
- If the errors in the independent variable are larger than those in the dependent variable then you may want to rewrite the expression thus interchanging the two quantities.

Goodness of Fit

A rule of thumb that some people use is that a good fit will pass within the error bars of about two-thirds of the data. In judging goodness of fit, there need both a quantitative measure and observation of trends and physical insight as qualitative guides. For the quantitative measure Equation (59) can be used as a basis.

If the values you determined for the slope and intercept give a line with a good fit, then you would expect evaluating the right-hand side of Equation (59) should give you a value close to n However, it is more than likely that you get a value different from n. The worse possibility is if the summation gives a value much greater than n.

$$n \gg \sum_{i=1}^{n} \left(\frac{y_i - f(x_i)}{\delta y_i} \right)^2 \tag{64}$$

There could some underlying reasons why the fit isn't as good as it could be.

1. You have defined the wrong function $f(x)$ for the fit. In the case of using a computer to calculate the fit, check the commands you are giving the program.
2. You have used the wrong data for the fit. In the case of using a computer to calculate the fit, be sure you are giving the program the correct information.
3. Your uncertainties were underestimated. Review your uncertainties to check that you haven't assumed your data is more precise than it really is.

The other extreme possibility is if the summation gives a value much less than equation n.

$$n \ll \sum_{i=1}^{n} \left(\frac{y_i - f(x_i)}{\delta y_i} \right)^2 \tag{65}$$

This can result if your uncertainties are overestimated. If you do get values that are very different from n, then you should try to determine if there are any corrections that you can make until you have a value close to n.

References.

Here are a few books the topics in this document in more detail.

Non-technical

These titles are available through the Mansfield Library.

1. Darrel Huff, *How to Lie with Statistics*.

Technical

These titles are available through the Mansfield Library.

2. J. R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*.
3. P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*.
4. Hugh D. Young, *Statistical Treatment of Experimental Data*.

These titles may be more difficult to find, as the are not in the Mansfield Library.

5. C. M. Reidel, *Least Squares Fitting of Data*.
6. Norfolk State University, *Measurement Analysis 2: Probabilistic Uncertainty, Least Squares Fitting, and Graphical Analysis*.