Random Error in Data

Abstract

In this lab, you will measure reaction time by catching a dropped ruler and converting the distance fallen to a time. Sufficient data will be taken to determine whether the error in this experiment is normally distributed – i.e., random.

Introduction

The difference between science and not-science is data – conclusions in science are guided by whether or not they are consistent with actual measurements. For example, a reasonable question is whether all objects fall to the ground at the same speed (they don't) or same acceleration (they don't) or whether all objects falling near the surface of the Earth subject ONLY to the gravitational force exerted by the Earth fall with the same acceleration (surprisingly, they do). Aristotle asked the question “Will a heavier object fall to the ground before a lighter one?” and answered in the following way:

A heavier object feels a larger force than a lighter object (which is what we mean by “heavier”). Experience shows clearly that an object subject to a larger force speeds up faster and thus travels faster and will hit the ground first. Hence, heavier objects will hit the ground first.

Galileo countered with the following:

Suppose Aristotle is right. Then, if we join two objects together they should fall SLOWER since the lighter object will slow the heavier object down. OR, they will fall FASTER since the two objects together are an even HEAVIER object. This is a paradox and the only way out of it logically is to refute the premise – that heavy and light objects fall differently. Ergo, heavy and light objects fall together.

Galileo’s argument is clever, but in the end he proved his point not with clever logical arguments but by simply dropping two objects and seeing that they fell to the ground together. In short, he did experiments to see how the world behaved. It’s not hard to figure out why it took so long for people to agree that this was the right strategy – it’s hard (see, for example: http://www.theonion.com/article/national-science-foundation-science-hard-1405)

The easiest part of science is coming up with a host of potential explanations for any effect – the hard part is ensuring that the effect is real and choosing EXPERIMENTALLY among all the possible explanations. A problem Galileo faced when demonstrating his conclusions to others was called “moving the goal posts”. After Aristotle’s believers agreed that a heavy object would hit the ground WELL before a light object, Galileo would drop them. And they would land virtually together. But, finding that the “larger outstrips the smaller by two finger-breadths”, deniers then claim to have been vindicated when the larger objects hits first – they have “moved the goal posts” so that a failed experiment is now a success. The REAL issue is error – how do we make scientific progress if every measurement is subject to some level of error. That is the topic of today’s lab – how to quantify and address error.
**Random Error**

Random is a tricky concept – if I write down the sequence: \{1, 6, 8, 3, 4, 4, 3, 5, 8\} and ask you the next few digits, I think you would have a hard time coming up with \{7, 3, 7\}. But if I told you that I was writing every fourth digit of pi, I think you could do it easily (after a quick Google search :). And when it comes right down to it, probably NOTHING is TRULY random – when you flip a coin or roll a die, the final state (heads, say, or a number 1) is probably predetermined by the initial conditions when you toss them and the laws of physics. But we can STILL consider them “random” since there is no PRACTICAL way to predict the final state. Similarly, in an experiment involving the time for a dropped object to fall (like the one we will do today), the error caused by a person slamming a door upstairs or a surge in the AC voltage in the room can PROBABLY be predicted IN PRINCIPLE, but in practice they are random errors. And random in that they can cause either an increase OR a decrease in the measured time (i.e., cause the measured time to be either too long or too short). So we'll consider the effect of random error on the time for an object to fall a given distance and draw whatever general conclusions we can. First, the practice of science assumes that there IS a result, somewhere out there in the Universe – a “correct” answer. And that an experiment is likely to get a different answer because of various sources of error (not just the two I've suggested, but a great many). These errors are not likely to be the same (say 1%) but are likely to be different. But it turns out that we can find out what random error looks like by considering exactly that unrealistic case – a small number of error sources (in this example, five) each causing the same error (1%) in every experiment, but random in direction (±1%). The final assumption (frequently made) is that the errors in question are “uncorrelated” - that is, that the error made by the door slam is not affected by the error caused by the power fluctuation. So with all these presumptions about the error, what would an experiment look like?

**The Random Walk**

First, in this unrealistic case, there are only five possible results for each experiment. Suppose that you are unlucky and ALL five errors are +1%. (For later, write this as the \[+,+,+,+,+\] result – all 5 errors are positive). Then the experimental result will be 5% above the “correct” result. Similarly, you could be unlucky and have all errors be negative (write this as \[-,-,-,-,-\]) giving a result 5% BELOW the “correct” result. (We will call these extreme results “outliers” - we'll see how to identify outliers later). But note that another possibility is that error #1 is 1% LOW while the other four are all high (write this as \[-,+,+,+,+\]). In this case, the error is +3% ; that is, 3% above the “correct” answer. But note that there are other ways to get an error of +3%. In fact, for so few sources of error, I can write out ALL the ways to get +3%: \{[-,+,+,+,+], [+,-,+,+,+], [+,+,+,+,+]\} and since I can't see each individual error, these all look the same to the experiment: 3% too high. So there's only one way to get +5%, but FIVE ways to get +3%. So which result is more likely? (HINT: +3%). And if error #1 AND #3 are both -1%, \[-,+,-,+,+\], the error is +1%. If you write them all out (and you should), you find there are 10 possible ways of getting a +1% error. Since the results are the same for the net negative error (one way to get -5%, five ways to get -3% and 10 ways to get -1%), you get the following if you plot the possibility of getting a certain error as a function of that error:
It can be seen that a 5% error is possible, but less likely, since ALL errors must ALL be of the same sign (Positive or negative). Results with such large variations from most data are called “outliers”. They happen, but not often (if they happen often, they aren’t outliers). This analysis becomes more realistic the greater the number of sources of error you consider (few experiments have ONLY 5 sources of error). But explicitly counting up all the possibilities for 10 errors, for example, becomes unjustifiably tedious. Fortunately, the mathematical operation called Combination (see Wikipedia) helps here. Using that, the plot of 10 sources of error looks like:

and for 100 sources of error looks like:
There are a few things to notice. One, the distribution becomes smoother the more sources of error there are. Two, the distributions widens as the number of sources of error increases. That is, the more sources of error, the greater the overall error in measurements. But note that it grows slowly. For the first case (five sources of error) most experimental results are within about ±2% of the center. For 10 sources, the results are within ±3% and for 100 errors, within ±10%. That is, the error grows as the square root of the number of sources of error: \( \sqrt{N} \). Note that the AVERAGE error is found by dividing by the number of measurements, \( N \). So the FRACTIONAL error will be proportional to: \( \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \). There are two points here – 1) the fractional (or percent) error DECREASES the more data you take (which makes sense) and – 2) there is NO point at which you have enough data. The fractional error ALWAYS decreases if you take more data. And it’s always hard – taking 10 times as much data always decreases the error by a factor of \( \sqrt{10} \approx 3 \) (welcome to science). So know now that whenever you ask me how much data you should take, the answer will always be “more”.

Ultimately, how much data you take is limited by the time and interest you have.

Finally, the shape of the error distribution may look familiar. It's called a normal distribution or a Gaussian function (named for the mathematician Johann Carl Friedrich Gauss – see Wikipedia). So we see why a random error causes a peaked distribution of errors – only in the unluckiest of events do ALL the errors go in one direction. Today's lab will be about determining experimentally whether your reaction time, measured as the time to catch a dropped ruler, has a random error and thus a Gaussian distribution of times. One last point. If you're a physicist, a chemist or a mathematician, probably the most important part of this lab to understand is the model of the random walk (see Wikipedia) – it's probably the second most important model in physics (after the damped, driven harmonic oscillator).

**The Lab**

The lab procedure couldn't be simpler. You will measure your reaction time by having a colleague suspend a ruler just above your thumb and finger. She will drop the ruler and you will catch it. Record the mark at which your finger caught the ruler. We will analyze this data as well as the reaction time we calculate later. Here is some example data taken by yours truly:

<table>
<thead>
<tr>
<th>D [inches]</th>
<th>6.5</th>
<th>6.5</th>
<th>6.19</th>
<th>6.56</th>
<th>8.38</th>
<th>6.75</th>
<th>7</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.19</td>
<td>6.5</td>
<td>8.5</td>
<td>5.38</td>
<td>7.13</td>
<td>5.63</td>
<td>3.63</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7.44</td>
<td>4.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Unfortunately for me, the ruler I had was marked in inches so I'll have to convert the units. Also, as we'll see, this is really not enough data to address the questions we care about.

**The Analysis**

The first question is which of these is the right answer? As your textbook ("Introduction to Error Analysis", John R. Taylor) shows, the best estimate of the “correct” answer, given a number of measurements, is to average all results. THIS IS TRUE ASSUMING THE ERROR IS RANDOM – I.E., NORMALLY DISTRIBUTED. In the vast majority of cases, people simply assume this to be true (as will we through most of this course). In today's lab, we will take enough data to actually confirm or deny this assumption (or at least to argue it's plausibility). If I calculate this average, the result is: 0.1600 meters. Next, to quantify the error in this measurement, your text directs you to calculate the standard deviation (the Wikipedia article is excessively detailed, but shows the connection to the normal distribution). The result is: 0.0276 meters. You should confirm each of these results (by hand AND with your calculator). To confirm that the error is normally distributed takes more effort. We have to do something similar to what was described in the Introduction. Unfortunately, we can't know how many sources of error there are. So we just guess. I used 10, you can try more or less and see what happens. So I divided the distance above into 10 intervals, starting at the minimum value (0.092 m) and ending at the maximum value (0.22 m). The intervals are then:

| 0.09 | 0.1 | 0.12 | 0.13 | 0.14 |
(Note that there are 11 entries). I then calculated how many of the results occur in each interval. From example, there is only one experimental result in the first interval (0.092075 is between 0.09 and 0.1045). There are two results between 0.14 and 0.155. And so on. The result looks like:

<table>
<thead>
<tr>
<th>Interval</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.09</td>
<td>1</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>0.12</td>
<td>1</td>
</tr>
<tr>
<td>0.13</td>
<td>1</td>
</tr>
<tr>
<td>0.14</td>
<td>2</td>
</tr>
<tr>
<td>0.15</td>
<td>3</td>
</tr>
<tr>
<td>0.17</td>
<td>8</td>
</tr>
<tr>
<td>0.18</td>
<td>5</td>
</tr>
<tr>
<td>0.19</td>
<td>3</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>0.22</td>
<td>1</td>
</tr>
<tr>
<td>0.22</td>
<td>1</td>
</tr>
</tbody>
</table>

You should do this calculation by hand so that you understand what's going on. But I have an OpenOffice spreadsheet posted on the Lab website that you can download and adapt for the more tedious analysis you will have to do with your data. (You will have to learn a little bit about OpenOffice. If you don't want to do that, you can use Excel or whatever else you want, but I probably won't be able to help you as much. Dr. Allain would say you should use Vpython/Glowscript – he's right, but again, I won't be much help. If you're “not good with numbers”, I can help with that. If your plan is to remain “not good with numbers”, you're likely to have trouble with this course.) If I now graph this distribution, it looks like:

**Probability distribution for Dropped Ruler**

![Graph of Probability distribution for Dropped Ruler](image)

I've included a Gaussian fit that is mostly for me, so don't worry about it (but if you're interested, ask...
me about it and I’ll explain what was done). You can see that the idea that this data is random (normally distributed) is not crazy, but the there’s really not enough data to be really confident. But note the peak of the distribution (about 0.15 m) is pretty close to the average calculated earlier (0.16 m). And the width of the distribution (0.025 m) is close to the standard deviation we calculated earlier (0.0276 m). This is to be expected IF the error is random.

But the real question we want to answer is the reaction TIME, not the distance the ruler falls. Do do that, we need to relate the time taken to fall to the distance fallen. As you will learn this semester, that relationship is:

\[ D = \frac{1}{2} a t^2 \]  

(1)

and so the reaction time is:

\[ t = \sqrt{\frac{2D}{a}} \]  

(2)

In this case, the acceleration is the acceleration produced by the interaction of the Earth and the ruler (the ruler’s “weight”). As we’ll discuss further, that acceleration is expected to be constant at 9.81 m/s².

So, if the ruler falls 0.216 m before I catch it, then it fell for 0.210 s. And so my reaction time was for that drop was 0.210 seconds. Doing this for all the data, I get an average reaction time of 0.180 s, with a standard deviation of 0.016 s. That is, this data tells me that my best estimate for the “correct” answer to “what is my reaction time?” is 0.180 s. But there is error (from whatever sources) of 0.016 s. So, I then create another distribution, this time for the reaction time. The final result is:

![Probability of reaction time](image)

Fitting Model:

\[ y = a + b \times \exp(-(x-c) \times (x-c)/(d \times d)) \]

- a = 0.0266517182
- b = 0.284996927
- c = 0.185446492
- d = 0.0096312503
So, according to this result, the best estimate for my reaction time is 0.185 s, with a likely error of 0.010 s. (Not particularly fast :).

**Conclusions**

Obviously, the mechanics of this lab are short and simple (see, for example, this lab designed at a much lower level: [https://backyardbrains.com/experiments/reactiontime](https://backyardbrains.com/experiments/reactiontime))

The real difficulty is in the analysis and the underlying ideas that I hope you will come to understand. Get used to it. It's how I roll. But just one particular measurement of reaction time (and the data above is ONE measurement of reaction time) is not particularly interesting. Rather, it's more meaningful to see how various conditions affect the reaction time. For example, is your reaction time different for right hand or left hand? Is there a similar difference (if any) for a left handed person? Is there a difference between visual reaction (you just see the ruler drop), auditory reaction (you close your eyes and your colleague gives a verbal cue when they drop the ruler) and both (you see the ruler drop AND hear the cue)? Surprise me.