Systematic Uncertainty

Max Bean

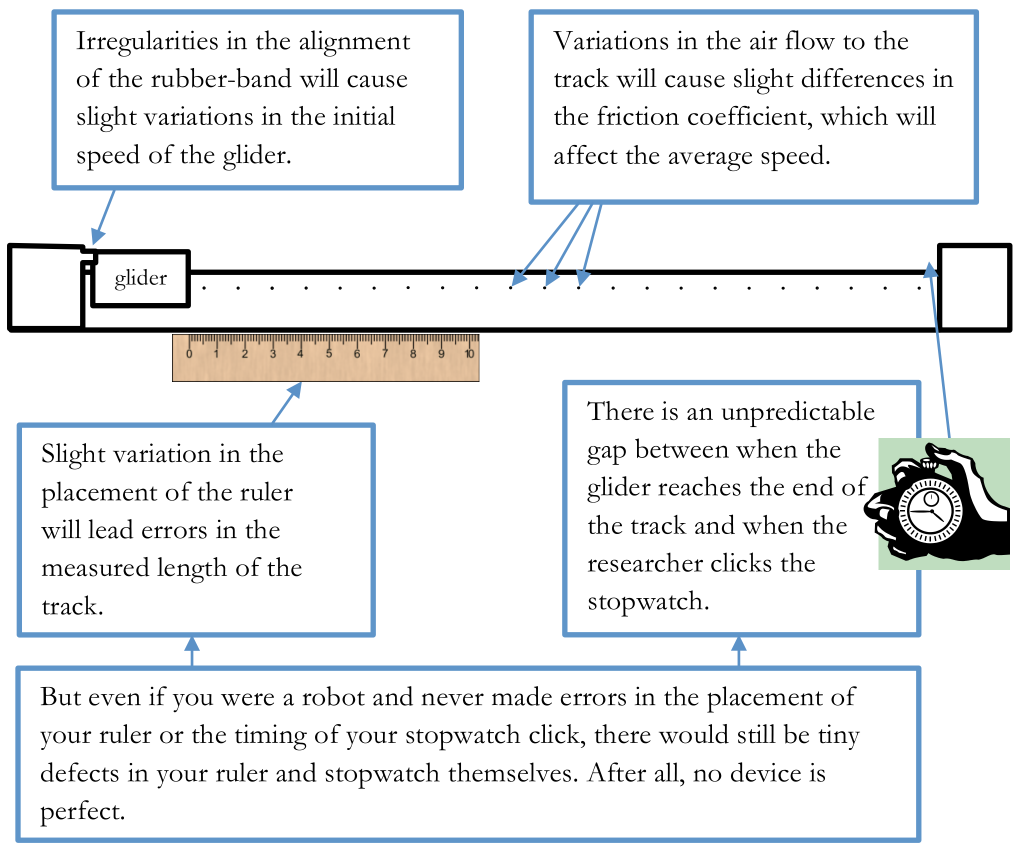
John Jay College of Criminal Justice, Physics Program

When we perform an experiment, there are several reasons why the data we collect will tend to differ from the actual values we are trying to measure. Say, for example, that you're measuring the average speed of a glider moving down an air-track. You will have at least three types of RANDOM variation:

Random variation in the actual thing you're trying to measure:

Random human measurement error:

Random defects in your measurement devices:



THE BAD NEWS: because these types of error are RANDOM, they are very difficult to quantify. You can never say exactly how big the effects of random errors are. Statistics allows us to say things about the PROBABILITY that these error effects are a given size, but it can never say anything with 100% certainty.

THE GOOD NEWS: the fact that these errors are RANDOM means that we can reduce their effect by AVERAGING MULTIPLE TRIALS. Again, we can never eliminate them completely, but the more trials we run, the smaller they get. We can also do cool stuff with scatter plots to help us deal with random variation; we'll do some of that later on.

THE OTHER GOOD NEWS: in this course, we are not really that interested in RANDOM errors, and you are NOT REQUIRED TO ESTIMATE THEIR EFFECTS in your lab reports (which is good news for you, because statistics is a pain). What we are definitely interested in is something called SYSTEMATIC UNCERTAINTY, and you will definitely have to estimate the effect of this in your lab reports.

Systematic uncertainty is a QUANTIFIABLE uncertainty caused by the very nature of measurement. Let's go back to our glider example. Imagine for a moment that the glider & the track are perfect: the glider always goes down the track at EXACTLY the same speed. (That's impossible in the real world, of course, but in physics we deal with an imaginary perfect world that's *similar* to the real world, but *not the same*.) Imagine, also, that you are a perfect robot researcher: you always click the stop watch at the EXACT MOMENT that the glider reaches the end of the track; you always line up your ruler PERFECTLY with the end of the glider; etc. (Also impossible.) Finally, imagine that you have perfect measurement devices with no defects. (Impossible again.) Even in this imaginary, perfect scenario, there would STILL be uncertainty in your data. Here's why:

Say you use your perfect robot hand to line up the meter-stick perfectly, and with your perfect robot eyes, you see that the distance from the end of the glider to the end of the track comes to 572 mm. But wait: is the track really a *perfectly* whole number of millimeters long? Isn't it really a few fractions of a millimeter over or under? It's not that your meter-stick is defective; it's just that it doesn't have fine enough gradations. Of course, we could get a special measuring device with microscopic gradations & read it using a magnifying glass or lasers or something, but whatever device we use MUST have a SMALLEST UNIT, and VARIATIONS SMALLER THAN THAT UNIT WILL BE & SHOULD BE UNDETECTABLE.

This is NOT an ERROR; it is an intrinsic part of what it means to measure things; it is a **SYSTEMATIC UNCERTAINTY. When we measure things, we don't get VALUES, we get RANGES OF VALUES.**

When we measure the glider track with our meter-stick and say that it's 572 mm long, we don't mean that it's EXACTLY 572 mm; we mean it's CLOSER to 572 mm than it is to 571 mm or 573 mm. In other words, it's between 571.5 and 572.5; we write it like this: 572 mm $\pm$0.5 mm. IN GENERAL: a measurement taken on a measuring device is DEFINED to be off by up to one half of the smallest unit captured by the measuring device. If we're measuring with a ruler whose smallest units are eights of an inch, our uncertainty will be $\pm$ half of an eighth of an inch, i.e. $\pm \frac{1}{16}$ inch. This $\pm$ value is called the ABSOLUTE UNCERTAINTY INTERVAL, but we will mostly refer to it as simply the uncertainty interval in this packet. Notice that the absolute uncertainty interval HAS ITS OWN UNITS.

(One quick side note: you have to be careful with digital measuring devices like stopwatches. Ignoring hundredths of a second, a stopwatch typically shows $\boxed{\text{0:00}}$ until the very end of the 1st second, then shows $\boxed{\text{0:01}}$ until the end of the 2nd second, and so on. So, when your stopwatch is showing $\boxed{\text{0:25}}$, that means that it has been running for anywhere from 25 to just under 26 seconds. Therefore, we should record our measurement as 25**.5** seconds $\pm$ 0.5 seconds.)

That's all you need to know about uncertainty for RAW DATA, i.e. **MEASURED** QUANTITIES. But once you've collected your raw data, you'll have to ANALYZE it by adding, subtracting, multiplying, & dividing measurements, in order to obtain **ANALYZED** QUANTITIES, which will have their own uncertainty intervals. The rest of this packet is about how to figure out the uncertainty intervals of analyzed quantities. Everything in this packet is based on The Basic Question Regarding Uncertainty Intervals, which is…

"What are the maximum and minimum possible values of the quantity?"

In our meter-stick example above, the MAXIMUM POSSIBLE VALUE was 572.5 mm. The MINIMUM POSSIBLE VALUE was 271.5 mm. Note what we're NOT asking here: we're NOT asking which value or range of values is most PROBABLE. We're only interested in what's POSSIBLE. This is good news, because, among other things, it makes the math much simpler.

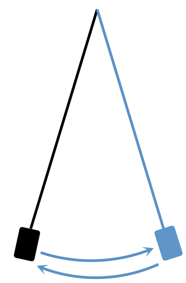
Mathematical Methods for Calculating Uncertainty Intervals

What follows are four different methods for calculating uncertainty intervals in different situations. We will provide the specific mathematical procedures for each case, and we'll explain those procedures, but there's nothing magical happening with these procedures; they all come naturally from the central question regarding uncertainty intervals: What are the maximum & minimum possible values?

Method #1: Multiplying & Dividing by a Dimensionless Number

Sometimes, we have to divide our data by a dimensionless number (i.e. a number without units). For example, say we are trying to measure the period of a pendulum (see box at right). The period is pretty short—less than one second—so instead of trying to measure a single trip back and forth, we count off 20 trips back and forth and time that. Then we divide our result by 20 to get the length of one period. In this case, 20 is a DIMENSIONLESS NUMBER. It does not correspond to any measurement. It has no uncertainty of its own and no units.

The **period** of a pendulum is the length of time required for the pendulum to complete one trip back and forth.



When dividing by a dimensionless number, divide the uncertainty range by the same number.
When multiplying by a dimensionless number, multiply the uncertainty range by the same number.

Example 1.A: In the pendulum example above, say we measured the length of 20 periods of our pendulum to be 15 seconds ± 0.5 seconds. We then divide by 20 to get the length of one cycle, and we get 0.75 seconds. We also divide our uncertainty range by 20: ± 0.5/20 = ±0.025. So our measurement for the length of one cycle would be 0.75 seconds ± 0.025 seconds.

Example 1.B: Say we're doing an experiment with Legos. We measure the height of a leg and find that it's 18mm ±0.5mm. Now we want to calculate the height of a tower that 50 Legos high. We multiply the height and get 18mm × 50 = 900mm. We multiply the uncertainty and get ±0.5mm × 50 = ±25mm. So our result is 900mm ±25mm.

One thing we can notice immediately from these examples is that measuring a larger quantity and dividing reduces uncertainty, while measuring a smaller quantity and multiplying increases uncertainty. You will want to keep this in mind when designing experimental procedures.

Proof of Method 1

You may think that Method 1 is so straightforward that it needs no proof. It's true that this method is very intuitive, but in some of the other methods in this packet some very un-intuitive things are going to happen (but you'll be able to see why), so the mathematically curious student may want to know how we can be so sure that what happens in Method 1 *is* so straightforward and intuitive. Here's the proof, for those who are curious:

Remember that, when developing our methods for operating on uncertainties, we're always interested in one basic question: "What are the maximum & minimum possible values of the quantity?"

With that in mind, let's begin with an example. In Example #1 above, we measured the duration of 20 periods on our pendulum and found that they took 15 seconds ± 0.5 seconds altogether. The **maximum**

possible value of the measured quantity is 15.5 seconds. The **minimum possible value** of the measured quantity is 14.5 seconds. We divide each of these by 20:

15.5/20 = 0.775 = Maximum possible length of one period

14.5/20 = 0.725 = Minimum possible length of one period

So, our calculated value for the period of our pendulum is **0.75 ± 0.025**

Now, here's the true algebraic proof, with variables. Given: a measured value m with an uncertainty interval u, and a dimensionless number k, find the uncertainty interval of (m/k) and $(m \times k)$. The maximum possible value of the quantity is $(m + u)$. The minimum possible value is $(m - u)$. So, we have

Maximum value of quantity divided by $k = \frac{m+u}{k} = \frac{m}{k} + \frac{m}{u}$

Minimum value of quantity divided by $k = \frac{m-u}{k} = \frac{m}{k} - \frac{m}{u}$

Maximum value of quantity multiplied by $k = k(m + u) = km + ku$

Minimum value of quantity multiplied by $k = k(m - u) = km - ku$

Thus, m/k has an uncertainty interval of $\pm\ u/k$,
and km has an uncertainty interval of $\pm\ ku$

Method #2: Adding & Subtracting Measurements

Say that we know the total length of the air track in our glider lab and we know the length of our glider. We can find the distance traveled by the glider by subtracting the glider length from the track length. Or say we now want to know the mass of the glider with a lab rat riding on top of it. We know the mass of the glider and we know the mass of the rat, so we just add them to find their combined mass.

In these cases, we are adding or subtracting two measured values, each of which will have its own error interval. One thing to notice in passing: if you are adding or subtracting measured values (*unlike* when you're multiplying and dividing), THEY MUST BE IN THE SAME UNITS. This means that they were probably gathered using the same measuring device, which means they ought to have THE SAME UNCERTAINTY INTERVAL. (This is also why you will **never** have to add or subtract a dimensionless number to/from a quantity.)

So, here are the rules:

When ADDING measured quantities, you ADD the uncertainty intervals.
When SUBTRACTING measured quantities, you <u>still</u> ADD the uncertainty intervals.

Example #2.A: In the glider-riding lab rat example, above, say we find that:

Mass of glider = 0.148 kg ± 0.0005 kg.

Mass of lab rat = 34g ± 0.5g—or 0.034 kg ± 0.0005 kg.

If we add these to find the total mass of the glider with the rat riding on it, we get:

(0.148 kg + 0.034 kg) ± (0.0005 kg + 0.0005 kg) = 0.182 kg ± **0.001** kg

Example #2.B: Say we found that:

our air track is 684 mm ± 0.5 mm long and

our glider is 112mm ± 0.5 mm long.

If we subtract these to find the distance traveled by the glider, we get:

(684mm – 112mm) ± (0.5mm + 0.5mm) = 572mm ± **1.0**mm

The first rule for method 2 (the rule for addition) is pretty intuitive. The second rule (the rule for subtraction) is not. To see why it works, we'll go back to (you guessed it!) our basic question about uncertainty intervals: what are the maximum and minimum possible values of the quantity?

In example 2B above, we have:

length of air track: L_a = 684 mm ± 0.5 mm; so L_a maximum = 684.5 mm and L_a minimum = 683.5 mm
length of glider: L_g = 112 mm ± 0.5 mm; so L_g maximum = 112.5 mm and L_g minimum = 111.5 mm

The MAXIMUM POSSIBLE VALUE of $L_a - L_g$ will be when L_a is BIG and L_g is SMALL. Thus,

$$L_a - L_g \text{ maximum} = 684.5 \text{ mm} - 111.5 \text{ mm} = 573 \text{ mm}$$

The MINIMUM POSSIBLE VALUE of $L_a - L_g$ will be when L_a is SMALL and L_g is BIG. Thus,

$$L_a - L_g \text{ minimum} = 683.5 \text{ mm} - 112.5 \text{ mm} = 571 \text{ mm}$$

So, $L_a - L_g$ can be anywhere from 571 mm to 573 mm, i.e. **572 mm ± 1 mm**
Because both quantities could be off in **either direction**, the total uncertainty of the difference is the sum of the uncertainties.

Notice that, because of this pair of rules,

- measuring a difference directly creates lower uncertainty than measuring two quantities and subtracting; and
- measuring a combined quantity directly creates lower uncertainty than measuring two quantities & adding them.

BUT these effects are small and may be offset by error factors; for example, it may be easier to line up the meter stick precisely to measure the length of the glider and the length of the track than it is to line it up to directly measure the length of the distance from the end of the glider to the end of the track.

When designing experimental methods, you should take into consideration BOTH factors that increase/decrease **uncertainty** AND factors that tend to increase/decrease **error**.

Method #3: Averaging Multiple Trials

One interesting result of thinking about UNCERTAINTY instead of ERROR and worrying about POSSIBLITY instead of PROBABILITY is that AVERAGING MULTIPLE TRIALS DOES NOT REDUCE UNCERTAINTY.

(But, as we saw in method #1, measuring a larger quantity and dividing *does* reduce uncertainty.) Instead,

When averaging multiple trials, you average the uncertainty interval from each trial.

Example 3.A: Say I'm measuring the speed of my glider; I perform 3 trials, and get the following readings for the gilder's time: 13 s ± 0.5 s, 14 s ± 0.5 s, and 12 s ± 0.5 s. When I average these, I will get:

(13 + 14 + 12)/3 s ± (0.5 + 0.5 + 0.5)/3 s = 13 s ± 0.5 s.

Note that, in the above example, because the things we were averaging were all the same kind of measurement, their uncertainty intervals were all the same, so averaging them left them the same. This will often be the case—but NOT ALWAYS. It will NOT be the case if the things we are averaging are calculated values, like velocity, that require multiplying or dividing.

Method #4: Multiplying & Dividing Measurements

In the glider lab that we've been discussing throughout this packet, we're ultimately trying to calculate the average speed of the glider. Average Speed = Distance/Time, so we're going to have divide our distance (a measured quantity) by our time (another measured quantity). Note that **this is completely different from dividing by a dimensionless number**: time has units and an uncertainty interval of its own. To take another example, if we were calculating an area or a volume, we would have to multiply two or three measured distances together, each of which would have units and uncertainty intervals.

The method for calculating uncertainty intervals for these operations is more complicated than the others, and it can't be put in a quick, one-line rule—but don't worry, it's not that hard. Just remember our basic question about uncertainty intervals: what are the maximum and minimum possible values of the quantity?

Ok, with this in mind, notice the following:

In **multiplication**, BIGGER×BIGGER=BIGGEST and SMALLER×SMALLER=SMALLEST

In **division**, BIGGER/SMALLER=BIGGEST and SMALLER/BIGGER=SMALLEST

Think about that for a moment and make sure it makes sense. Then we'll do some examples.

Example 4.A: Say we measured our glider's distance to be 571 mm ± 0.5 mm, and we measured its time to be 13.5 s ± 0.5 s. So, the **maximum possible value** for the speed will be

MAX DISTANCE/MIN TIME = 571.5mm/13s ≈ 44.0 mm/s

And the **minimum possible value** for the speed will be

MIN DISTANCE/MAX TIME = 570.5mm/14s ≈ 40.8 mm/s

Also, if we calculate the speed **ignoring uncertainty**, we get

572/13.5 =42.3 mm/s

Example 4.B: Say we are calculating the volume of a water tank, and we take the following measurements: height = 40 mm ± 0.5, length = 50 mm ± 0.5, and width = 30 mm ± 0.5. So, the **maximum possible value** for volume will be

MAX H × MAX L × MAX W = 40.5 mm × 50.5 mm × 30.5 mm = 62,380.125 mm^3

And the **minimum possible value** for volume will be:

MIN H × MIN L × MIN W = 39.5 mm × 49.5 mm × 29.5 mm = 57,679.875 mm^3

Also, if we calculate the speed **ignoring uncertainty**, we get

40 mm × 50 mm × 30 mm = 60,000.00 mm^3

Note that, in both examples above, **our calculated value is not exactly in the middle of our uncertainty interval!** This is an ASYMMETRICAL UNCERTAINTY INTERVAL, and it is a standard feature of our method for division and multiplication. There are three ways you can express it:

1. Simply provide the uncertainty interval with no specific data point:
   Average speed of glider is between 40.8 mm/s and 44.0 m/s
2. Present your calculated data point along with the asymmetric uncertainty interval:
   Average speed of glider = 42.3 mm/s, (Min: 40.8 m/s, Max: 44.0 m/s)
3. Calculate the midpoint of the new uncertainty interval and use this as your data point:
   Average speed of glider = 42.4 mm/s ± 1.6 mm/s

Any of these three formats is ok—in terms of scientific practice & in terms of the grade on your lab report—but DO please THINK about the advantages of each BEFORE selecting one.

Uncertainty Interval Calculation – a Side-by-Side Comparison

What we have presented in this packet is the new **Absolute Uncertainty Method** for calculating uncertainty intervals. Previously, JJay physics classes have used the **Fractional Uncertainty Method**. You are welcome to use either one in your lab reports. We're not going to go into detail about how the Fractional Uncertainty Method works, but you can find a quick overview of it in this side-by-side comparison of the two methods.

<table>
<tr><th>The Fractional Uncertainty Method</th><th>The Absolute Uncertainty Method</th></tr>
<tr><td colspan="2">Before You Begin To Analyze Your Data…</td></tr>
<tr><td>After collecting raw data, convert each uncertainty interval to a fractional uncertainty (FU) by dividing it by the measured value:

Examples A: Given 45m ± 0.5m
FU = 0.5m/45m = 0.011 or 1.1%

Example B : Given 28s ± 0.5s
FU = 0.5s/28s = 0.018 or 1.8%

(Notice that fractional uncertainty has NO UNITS!)

Now you're ready to analyze your data…</td><td>For the Absolute Uncertainty Method, you do not need to do any conversion before you begin to analyze your data. Your raw data comes with an absolute uncertainty interval (e.g. ± 0.5m or ± 0.0005kg) and you will use this interval to calculate all the uncertainty intervals for your analyzed quantities.</td></tr>
<tr><td colspan="2">When Multiplying/Dividing by a Dimensionless Number…</td></tr>
<tr><td>Fractional uncertainty remains unchanged

Example A:
45m (FU=0.011) × 10 = 450m (FU=0.011)</td><td>Multiply/divide the absolute uncertainty interval by the same number.

Example A: (45m ± 0.5m) × 10 = 450m ± 5m
Example B: (45g ± 0.5g) / 10 = 4.5g ± 0.05g</td></tr>
<tr><td colspan="2">When Adding/Subtracting Measured Quantities…</td></tr>
<tr><td>Use the largest fractional uncertainty from among the quantities that you are adding/subtracting.

Example:
23s (FU=0.014) + 31s (FU=0.012) = 54 (FU=?)
Use the larger fractional uncertainty: 0.014
Sum = 54s (FU=0.014)</td><td>Add the absolute uncertainty intervals.

Example A: (9s ± 0.5s) + (11s ± 0.5s) = 20s ± 1s
Example B: (30cm ± 0.5cm) – (16cm ± 0.5cm) =
14cm ± 1cm</td></tr>
<tr><td colspan="2">When Averaging Multiple Trials…</td></tr>
<tr><td>Use the largest fractional uncertainty from among the quantities that you are averaging:

Example:
Given three trials: 20s (FU=0.020), 25s (FU=0.018) = 30 (FU=0.017)
Use the larger fractional uncertainty: 0.020
Average = 25s (FU=0.020)</td><td>Average the uncertainty intervals…

Example:
Given three trials: 20s ± 0.5s, 25s ± 0.5s, 30s ± 0.5s
Average = 25s ± 0.5s</td></tr>
</table>

<table>
<tr><td colspan="2">When Multiplying/Dividing Measured Quantities…</td></tr>
<tr><td>Add the fractional uncertainties

Example A:
60mm (FU=0.008) × 70mm (FU=0.02) =
4200mm (FU=0.028)

Example B:
23mm (FU=0.014) / 7s (FU=0.025) =
3.3mm/s (FU=0.039)</td><td>Find the max possible values: big×big & big/small
Find the min possible values: small×small & small/big

Example A:
Given 20m ± 0.5m × 11m ± 0.5m:
Max Val = 20.5×11.5 = 235.75
Min Val = 19.5×10.5 = 204.75

Example B:
Given 34m ± 0.5m / 4s ± 0.5s:
Max Val = 34.5/3.5 = 9.86
Min Val = 33.5/4.5 = 7.44</td></tr>
<tr><td colspan="2">When You Have Completed Your Analysis and Found Your Final Quantities…</td></tr>
<tr><td>You must convert the fractional uncertainties <u>back</u> into absolute uncertainties for all outcome variables—i.e. quantities that you will discuss in your CONCLUSION. (But any variables that you will NOT discuss in your conclusion can be left with fractional uncertainties.) To do this, multiply the fractional uncertainty by the quantity it's associated with:

Example: 60g (FU=0.011) → 60g ± (60g×0.011) =
60g ± 0.66g</td><td>You're all done! Report the absolute uncertainty intervals that you calculated for each of your outcome variables.</td></tr>
</table>

Now that you've seen both methods, here are some practice problems. You can use EITHER METHOD to complete these.

You are trying to determine the density of a solid cylinder made of unknown material. You measure the cylinder and discover that it has a **diameter** of 25 mm ± 0.5 mm and a length of 45 mm ± 0.5 mm.

a) Use the formula $V = \pi r^2 h$ to determine the volume of the cylinder.

b) Calculate the ABSOLUTE UNCERTAINTY associated with the volume of the cylinder.
Hint #1: note that □ is a dimensionless number.
Hint #2: note that when you square r, you are multiplying r (a measured quantity) by r (a measured quantity).

c) You weigh the cylinder and discover that it has a mass of 79.422 kg ± 0.0005 kg. Use the formula $\rho = M/V$ (where ρ is density, M is mass, and V is volume) to find the density of the cylinder.

d) Calculate the ABSOLUTE UNCERTAINTY associated with the density of the cylinder.

And now, a conceptual question: why does it make sense that fractional uncertainty has no units?