Where do error analysis formulas come from?

Suppose we have an equation

$$x = \frac{1}{2}at^2$$

where x is position, a is acceleration, and t is time. Imagine the goal of some experiment is to measure both position and time for a number of data points and determine acceleration. At the end of the experiment, we want to express the final value of acceleration with appropriate sig figs.

One method is to take each data point and compute

$$a = \frac{2x}{t^2}$$

But how do small errors in the *measurements* of x or t affect the *calculation* of a?
First assume the measurement of t is a perfect measurement and only worry about errors in x.
Mathematically, this says assume t is constant and find the rate of change in a with respect to changes in x.

$$\frac{da}{dx} = \frac{2}{t^2}$$

Here da is uncertainty in the acceleration *calculation* while dx is uncertainty in the position *measurement*. We can rearrange this to see

$$da = \frac{2}{t^2}dx$$

The fractional error (similar to % error) in acceleration is $\frac{da}{a}$. If we divide each side of the equation by a we get

$$\frac{da}{a} = \frac{1}{a}\cdot\frac{2}{t^2}dx$$

$$\frac{da}{a} = \frac{1}{\frac{2x}{t^2}}\cdot\frac{2}{t^2}dx$$

$$\boldsymbol{\frac{da}{a} = \frac{dx}{x}}$$

Now repeat this process while holding x constant but allowing t to vary. We find

$$\frac{da}{dt} = -\frac{4x}{t^3}$$

$$da = -\frac{4x}{t^3}dt$$

$$\frac{da}{a} = \frac{1}{a}\left(-\frac{4x}{t^3}\right)dt$$

$$\frac{da}{a} = \frac{1}{\frac{2x}{t^2}}\left(-\frac{4x}{t^3}\right)dt$$

$$\boldsymbol{\frac{da}{a} = -2\frac{dt}{t}}$$

In real life, errors in both measurements could occur. We will assume errors from reading the meterstick are probably independent of errors in reading the stopwatch. This means I could be a little under or over in reading the meterstick; I could be a little over or under in reading the time. When this is true we add the errors in *quadrature*.

$$\boldsymbol{\frac{da}{a} = \sqrt{\left(\frac{dx}{x}\right)^2 + \left(2\frac{dt}{t}\right)^2}}$$

Notice the minus sign from the second bold equation (the error associated with time) doesn't matter.

Using propagated errors to estimate error in average value of a

Below is a data set I made up.

Things to notice:

1) The first few measurements of x have small errors (1 cm) while the last few have large errors (50 cm).
2) The sig figs for each value of x match the *column* of sig figs for the corresponding error dx.
3) The third column is the fractional error in x. Since it is an error calculation, I round to 1 sig fig. *Exception:* if the first digit of an error calculation is 1 then I include 2 digits.
4) The t, dt & $\frac{dt}{t}$ data come from time measurements and the sig figs were done the same way as the x, dx & $\frac{dx}{x}$ columns.
5) The acceleration was *calculated* (not *measured*) using $a = \frac{2x}{t^2}$. I will discuss sig figs in a second…
6) The da column was computed using $\frac{da}{a} = \sqrt{\left(\frac{dx}{x}\right)^2 + \left(2\frac{dt}{t}\right)^2}$.
7) Since the da column is an error calculation, I round to 1 sig fig. *Exception:* if the first digit of an error calculation is 1 then I include 2 digits.
8) I now know how many sig figs should appear on my acceleration *calculation*!

   I match the column sig figs for each a to the column of sig figs for the corresponding da.

x (m)	*dx* (m)	*dx/x*	*t* (s)	*dt* (s)	*dt/t*	*a* (m/s^2)	*da* (m/s^2)
1.42	0.01	0.007	1.0	0.1	0.10	2.8	0.6
5.34	0.01	0.0019	1.9	0.1	0.05	2.9	0.3
15.89	0.05	0.003	3.0	0.1	0.03	3.5	0.2
19.09	0.05	0.003	4.0	0.1	0.03	2.43	0.12
56.1	0.1	0.0018	5.0	0.1	0.02	4.43	0.18
62.9	0.1	0.0016	5.8	0.1	0.017	3.71	0.13
112.9	0.5	0.004	7.0	0.1	0.014	4.55	0.13
102.6	0.5	0.005	8.0	0.1	0.013	3.23	0.08

Since we expect each value of a to be the same in this case, we may average these values to get

$$a_{avg} = \bar{a} = 3.44499\,\frac{\mathrm{m}}{\mathrm{s}^2}$$

BUT HOW MANY SIG FIGS SHOULD I WRITE????

I estimate the *error* in the average by dividing the average error by $\sqrt{N}$ where $N = 8$ trials. This isn't a perfect estimate but it is not bad.

$$da_{avg} = d\bar{a} = \frac{\text{average of } da}{\sqrt{N}} = 0.0767\,\frac{\mathrm{m}}{\mathrm{s}^2} = 0.08\,\frac{\mathrm{m}}{\mathrm{s}^2}$$

I round to one sig fig because this is an error calculation.

I match the sig fig column for a_{avg} to the sig fig column of da_{avg}.

Our best estimate for the acceleration using *propagated* errors is

$$\boldsymbol{a_{avg} = 3.44 \pm 0.08\,\frac{\mathrm{m}}{\mathrm{s}^2}}$$

Getting the error using a graph

In this case, we assumed

$$x = \frac{1}{2}at^2$$

Compare this to the equation of a line

$$vertical\ coordinate = slope \times (horizontal\ coordinate) + intercept$$

If I allow $horizontal\ coordinate = t^2$ and $vertical\ coordiante = x$ I expect $slope = \frac{1}{2}a$.

Take a second to look at the equations and let this sink in…

I used the data table above to create the data table shown at right. Notice the following:

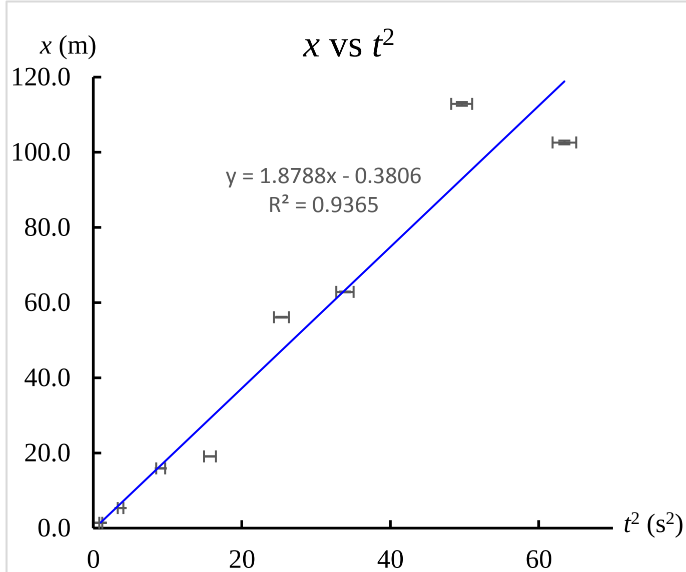
1) I calculated t^2 so the units also get squared.
2) If I *calculated* t^2, the error $d(t^2)$ must be *calculated* as we did for da. In this case $d(t^2) = 2t\ dt$.
3) After computing the error calculation of $d(t^2)$ I double checked the column of sig figs for each t^2 matches the sig fig column for the corresponding $d(t^2)$.
4) Notice the units of $d(t^2)$ are also s^2.

x (m)	dx (m)	t^2 (s^2)	$d(t^2)$ (s^2)
1.42	0.01	1.0	0.2
5.34	0.01	3.7	0.4
15.89	0.05	9.1	0.6
19.09	0.05	15.7	0.8
56.1	0.1	25.3	1.0
62.9	0.1	33.9	1.2
112.9	0.5	49.6	1.4
102.6	0.5	63.5	1.6

Plot on next page…

Here is the plot of $x\ vs.\ t^2$ corresponding to the data on the previous page. To notice:

1) Plots are named vertical coordinate vs horizontal coordinate…not the other way around.
2) Units do not appear in the plot title, only in the axis labels.
3) *Variables* are *italicized*, units are not.
4) The *horizontal* error bars correspond to $d(t^2)$.
5) The *vertical* error bars correspond to dx.
6) Notice how small the vertical error bars are. This makes sense. Look at how small the fractional errors $\frac{dx}{x}$ are compared to the fractional errors $\frac{dt}{t}$.
7) The slope is $1.8788\,\frac{\text{m}}{\text{s}^2}$. To get the units I used the units of rise (vertical axis) over units of run (horizontal axis).



According to our physics equations, we expected

$$slope = \frac{1}{2}a$$
$$a = 2 \cdot slope$$
$$a = 2\left(1.8788\,\frac{\text{m}}{\text{s}^2}\right)$$
$$a = 3.7576\,\frac{\text{m}}{\text{s}^2}$$

BUT HOW MANY SIG FIGS SHOULD I WRITE????

In this case I need

$$da = 2 \cdot d(slope)$$

I can determine the statistical error in the slope of a line using the LINEST function in Excel. This is discussed in depth with screen shots in the Lab Manual Appendices.

I used the function to determine $d(slope) = 0.199\,\frac{\text{m}}{\text{s}^2} = 0.2\,\frac{\text{m}}{\text{s}^2}$.

It is an error calculation…I round to 1 sig fig.

I plug this into the above equation to get

$$da = 2 \cdot d(slope) = 0.4\,\frac{\text{m}}{\text{s}^2}$$

I now match the column of error for my calculation of a to my calculated error da.

Our best estimate of a using statistical errors is

$$\boldsymbol{a = 3.8 \pm 0.4\,\frac{\text{m}}{\text{s}^2}}$$

Which procedure gives the best error estimate for a?

It depends. Generally speaking, I like to be as conservative as possible with my error estimates. By this I mean I tend to take the *larger* error estimate and go with that one. For this situation, I would choose to use the statistical error calculation from the slope and assume $a = 3.8 \pm 0.4 \frac{\mathrm{m}}{\mathrm{s}^2}$.

In some data sets, all the points lie nearly perfectly on the line. In those data sets statistical error is ridiculously small. When this happens, I tend to use the propagated error technique to determine my best estimate.

Why should I care?

For physicists, you need error analysis for sure in senior lab (three semesters of pure bliss).

For engineers, here is an article I found motivating why you might care. It told a good story to contextualize this.
https://www.designworldonline.com/why-its-important-to-always-use-tolerances/

Side note: if no tolerances are provided on a design specification a manufacturer would very likely estimate a tolerance using...

Wait for it…

THE COLUMN OF SIG FIGS YOU USED FOR YOUR NUMBERS.
This is why your instructors are always hounding you about sig figs.