Derivations for Error Propagation

James Pickering, School of Chemistry, University of Leicester

Disclaimer: This stuff is definitely not going to be examined and is here for interest (to satisfy the ~~nerds~~ *mathematically-inclined people*, like me).

The General Error Propagation Formula

Given a function $F = F(x, y, z, ...)$, we can define a thing called the **total differential**, $\mathrm{d}F$:

$$\mathrm{d}F = \left(\frac{\partial F}{\partial x}\right)\mathrm{d}x + \left(\frac{\partial F}{\partial y}\right)\mathrm{d}y + \left(\frac{\partial F}{\partial z}\right)\mathrm{d}z + ...$$

Where the partial derivatives in brackets are evaluated with the variables not in the derivative held constant, like normal. If we imagine we measuring x, y, or z, we could rewrite our $\mathrm{d}x$ as:

$$\mathrm{d}x = x_i - \bar{x}$$

provided that the measurement x_i was not very far from the mean $\bar{x}$, such that the infinitesimally small $\mathrm{d}x$ actually is infinitesimally small. We can also rewrite $\mathrm{d}F$ as:

$$\mathrm{d}F = F(x, y, z, ...) - F(\bar{x}, \bar{y}, \bar{z}, ...)$$

Putting all of that together gives us the horrific looking:

$$F(x, y, z, ...) - F(\bar{x}, \bar{y}, \bar{z}, ...) = \left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x}) + \left(\frac{\partial F}{\partial y}\right)(y_i - \bar{y}) + \left(\frac{\partial F}{\partial z}\right)(z_i - \bar{z}) + ...$$

We will leave this here for later, like some perfectly torn basil to throw into our mathematical pasta at a later point. Now let's consider if we can calculate the **mean of F**, that is $\bar{F}$. We can confidently say that:

$$\bar{F} = \frac{1}{N}\sum_{i=1}^{N} F(x_i, y_i, z_i, ...)$$

Just using the definition of the mean: adding up the values of F from every observation x_i, y_i and z_i, and dividing by the number of values. Let's now grab our mathematical basil and throw it in:

$$\bar{F} = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x}) + \left(\frac{\partial F}{\partial y}\right)(y_i - \bar{y}) + \left(\frac{\partial F}{\partial z}\right)(z_i - \bar{z}) + ... + F(\bar{x}, \bar{y}, \bar{z}, ...)\right]$$

All we've done here is take the horrific equation two lines above, and put it into the equation for the mean of F we just wrote above. All of those terms in the square brackets are being summed over, and we can just divide them up and do them one by one, because $\sum(a + b) = \sum a + \sum b$. Let's take the last one first:

$$\sum_{i=1}^{N} F(\bar{x}, \bar{y}, \bar{z}, ...) = F(\bar{x}, \bar{y}, \bar{z}, ...)\sum_{i=1}^{N}(1) = NF(\bar{x}, \bar{y}, \bar{z}, ...)$$

because our sum is over the variable i, and nothing in the $F(\bar{x}, \bar{y}, \bar{z}, ...)$ term depends on i, it is just a constant and we can pull it out of the front. Adding up the number 1 from $i = 1$ to N, just gives us N, so we end up with the result there. Now let's consider the other three terms:

$$\sum_{i=1}^{N} \left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x}) = \left(\frac{\partial F}{\partial x}\right)\sum_{i=1}^{N}(x_i - \bar{x}) = \left(\frac{\partial F}{\partial x}\right)(N\bar{x} - N\bar{x}) = 0$$

This orgasmic result arises because $\bar{x}$ doesn't depend on i, and from the definition of the mean $\bar{x}$[1]. You can also argue it from the idea that if x_i is randomly distributed around $\bar{x}$, if we take enough measurements, then $\sum(x_i - \bar{x}) \approx 0$. This generalises to the terms in y and z too, so we don't need to do those separately.

[1] Can you see where?

In the end, this means that only the final term in our big expansion is non-zero, and so:

$$\bar{F} = \frac{1}{N} N F(\bar{x}, \bar{y}, \bar{z}, ...) = F(\bar{x}, \bar{y}, \bar{z}, ...)$$

This is delicious. It means that the mean value of F can be determined by just measuring F, with all of its argument variables x, y, and z set to their mean values. We're going to use this shortly.

Let us now look the problem in the eye and write an expression for the error in F, or the standard deviation in F[2], ΔF. I'm going to write it in terms of the **variance** - the standard deviation squared - just because it tidies things up and removes the need for a big square root sign.

[2] These are functionally equivalent, because the standard error is directly proportional to the standard deviation, so the equations are the same regardless of what we call the error.

$$(\Delta F)^2 = \frac{1}{N-1} \sum_{i=1}^{N} [F(x_i, y_i, z_i, ...) - F(\bar{x}, \bar{y}, \bar{z}, ...)]^2$$

Where we have used our nice result from above (remember the basic standard deviation formula). We can now use what we did right at the top of the first page (the first equation I labelled 'horrific') and say:

$$(\Delta F)^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left[\left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x}) + \left(\frac{\partial F}{\partial y}\right)(y_i - \bar{y}) + \left(\frac{\partial F}{\partial z}\right)(z_i - \bar{z}) + ... \right]^2$$

This is horrible and we now have to expand the huge square bracket full of partial derivatives and things on the LHS. Let's not do this in all the gory detail because it's tedious and fairly easy, but we're clearly going to get a load of terms like:

$$\left(\left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x})\right)^2$$

And a load of cross terms like:

$$\left(\left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x})\right)\left(\left(\frac{\partial F}{\partial x}\right)(x_i - \bar{x})\right)$$

The trick is that **all of the cross terms go to zero** in the end, if the variables are independent and randomly distributed. These cross terms represent things called **covariances**, which tell you how one variable is dependent on another, but if they are all independent, all the covariances are zero by definition. This means that we can write our overall formula as:

$$(\Delta F)^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left[\left(\left(\frac{\partial F}{\partial x}\right)(x_i-\bar{x})\right)^2+\left(\left(\frac{\partial F}{\partial y}\right)(y_i-\bar{y})\right)^2+\left(\left(\frac{\partial F}{\partial z}\right)(z_i-\bar{z})\right)^2+...\right]$$

Alright. We are well on the way now, and can use our previous trick that $\sum(a+b)=\sum a+\sum b$ to split this thing up into individual x, y, and z parts. Let's look at what will happen to the x part and then generalise it to y and z:

$$\frac{1}{N-1}\sum_{i=1}^{N}\left(\left(\frac{\partial F}{\partial x}\right)(x_i-\bar{x})\right)^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{\partial F}{\partial x}\right)^2(x_i-\bar{x})^2$$

Now we can note that the partial derivative doesn't depend on i and so can be pulled outside the sum. Thus:

$$\left(\frac{\partial F}{\partial x}\right)^2\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2 = \left(\frac{\partial F}{\partial x}\right)^2(\Delta x)^2$$

Where we just use the definition of the standard deviation of x. We can now put everything back into the formula for $(\Delta F)^2$:

$$(\Delta F)^2 = \left(\frac{\partial F}{\partial x}\right)^2(\Delta x)^2+\left(\frac{\partial F}{\partial y}\right)^2(\Delta y)^2+\left(\frac{\partial F}{\partial y}\right)^2(\Delta y)^2+...$$

And arrive at the familiar formula for error propagation from the lecture notes! As you can see, the derivation is a bit of a slog. But it works quite nicely, I hope you agree. Well done if you made it this far (if you're reading this sentence and have followed it all, come and find me in my office for a reward).

The Relative Errors Formula

This derivation is mercifully shorter. We are going to show that for a function that is a product of a load of other things F:

$$F = a^p b^q c^r...$$

That the **relative error** $^{\Delta F}/_F$ is given by:

$$\left(\frac{\Delta F}{F}\right)^2 = p^2\left(\frac{\Delta a}{a}\right)^2+q^2\left(\frac{\Delta b}{b}\right)^2+r^2\left(\frac{\Delta c}{c}\right)^2+...$$

We will start with the formula we derived in the section above, which means we have to take partial derivatives of our expression for F, you can see that:

$$\left(\frac{\partial F}{\partial a}\right) = pa^{p-1}b^q c^r ...$$

and similar for the other derivatives, just using the normal rules of differentiating a polynomial, and the rules of partial derivatives (remember that if we are differentiating with respect to a, then the terms in b and c are constant). Thus, we can say:

$$(\Delta F)^2 = (pa^{p-1}b^q c^r ...)^2(\Delta a)^2 + (qb^{q-1}a^p c^r ...)^2(\Delta b)^2 + (rc^{r-1}a^p b^q ...)^2(\Delta c)^2$$

Now, we divide it all by F^2, lets just do the first term to keep it neat (and it all generalises in the end anyway). Replacing the second and third terms temporarily with a '...', we get:

$$\left(\frac{\Delta F}{F}\right)^2 = \left(\frac{pa^{p-1}b^q c^r ...}{F}\right)^2 (\Delta a)^2 + ...$$

Substitute in the definition of F, and we find:

$$\left(\frac{\Delta F}{F}\right)^2 = \left(\frac{pa^{p-1}b^q c^r ...}{a^p b^q c^r ...}\right)^2 (\Delta a)^2 + ...$$

Then everything nicely cancels to give:

$$\left(\frac{\Delta F}{F}\right)^2 = \left(\frac{p}{a}\right)^2 (\Delta a)^2 + ...$$

Let's put the other two terms back in, which behave in exactly the same way:

$$\left(\frac{\Delta F}{F}\right)^2 = \left(\frac{p}{a}\right)^2 (\Delta a)^2 + \left(\frac{q}{b}\right)^2 (\Delta a)^2 + \left(\frac{r}{c}\right)^2 (\Delta a)^2 + ...$$

Now, the whole thing simply rearranges to the result we wanted to show:

$$\left(\frac{\Delta F}{F}\right)^2 = p^2\left(\frac{\Delta a}{a}\right)^2 + q^2\left(\frac{\Delta b}{b}\right)^2 + r^2\left(\frac{\Delta c}{c}\right)^2 + ...$$

...Nice.