A Brief Note about Error Propagation

Because many relations in astrophysics are scaling relations or use logarithmic units like magnitudes, the agony of propagating errors is greatly reduced, and the computations made much simpler, if you work in terms of the relative uncertainties, σ_x/x. Relative uncertainties should already be familiar to you; it is what people have in mind when they say they've "measured something to 10%".

As an example, consider the following problem. We want to estimate the total mass of a visual binary system using Kepler's Third Law, which for this case is written in the form:

$$M_{tot} = (M_1 + M_2) = \frac{(a'')^3}{\varpi^3 P^2}$$

Here the "observables" are the angular semi-major axis, a'', the parallax, ϖ, and the orbital period, P. Each observable has an associated measurement uncertainty, $\sigma_{a''}$, σ_ϖ, and σ_P, respectively. The usual formula for formally estimating the absolute uncertainty on the derived total mass (M_{tot}), assuming all errors on the observables are uncorrelated, is:

$$\sigma^2_{M_{tot}} = \sigma^2_{a''}\left(\frac{\partial M_{tot}}{\partial a''}\right)^2 + \sigma^2_\varpi\left(\frac{\partial M_{tot}}{\partial \varpi}\right)^2 + \sigma^2_P\left(\frac{\partial M_{tot}}{\partial P}\right)^2$$

If you were to work this out algebraically you'd get an awful mess to evaluate numerically, raising the risks of making mistakes.

However, if you were to instead divide through by M_{tot} and work through the algebra, you will get the following much simpler formula for the *relative* mass uncertainty, σ_M/M:

$$\left(\frac{\sigma_{M_{tot}}}{M_{tot}}\right)^2 = \left(3\frac{\sigma_{a''}}{a''}\right)^2 + \left(3\frac{\sigma_\varpi}{\varpi}\right)^2 + \left(2\frac{\sigma_P}{P}\right)^2$$

This is a lot simpler to evaluate, since written this way you are just propagating the *relative errors* as the sum of their squares, with all of the contributions of the observables to the mass uncertainty neatly separated out and weighted by their respective powers! However, to get this formula took a lot of messy algebra... or did it?

Here's a dirty secret: I didn't to show all the intervening algebra above because I didn't do any algebra the first place…

Let's step back and look at the problem a little differently. Let's re-write the formula for the total mass in terms of *logarithms* of M_{tot} and the observables:

$$\begin{aligned}\log M_{tot} &= \log\left(\frac{a''^3}{\varpi^3 P^2}\right)\\ &= 3\log a'' - 3\log\varpi - 2\log P\end{aligned}$$

What is the uncertainty on $\log(M_{tot})$? From the error propagation, it is just

$$\sigma^2_{\log M_{tot}} = \left(3\sigma_{\log a''}\right)^2 + \left(3\sigma_{\log \varpi}\right)^2 + \left(2\sigma_{\log P}\right)^2$$

But, what is $\sigma_{\log(x)}$?

If y=ln(x), and x has uncertainty σ_x, then the uncertainty on y from the propagation formula above is

$$\sigma_y = \frac{\sigma_x}{x}$$

If y=log(x) (i.e., log base 10), it would be

$$\sigma_y = \frac{1}{\ln 10}\frac{\sigma_x}{x} \approx 0.4343\frac{\sigma_x}{x}$$

[In general, $\log_b(x) = \ln(x)/\ln(b)$]. Substituting this into the error formula from above, we get:

$$\left(\frac{\sigma_{M_{tot}}}{M_{tot}}\right)^2 = \left(3\frac{\sigma_{a''}}{a''}\right)^2 + \left(3\frac{\sigma_{\varpi}}{\varpi}\right)^2 + \left(2\frac{\sigma_P}{P}\right)^2$$

and note that the constant factor (1/ln10) has conveniently divided out, so I don't really need to make the distinction between $\log_{10}()$ and ln() in this case (Beware! *This is not always true*)! The coefficients of the relative errors on the right-hand side of the equation above are just the powers of the observables (i.e., 3 for a" and ϖ because they appear as the cube, 2 for P because it appears as the square). This is the secret to propagating relative uncertainties without recourse to truly ugly algebraic manipulation (and the risk of mistakes that introduces).

In general, you can always write down the correct formula for the propagation of relative uncertainties by looking at the powers on the observables. For example, if you are given this formula for z in terms of observables, a, b, and x:

$$z = a\frac{b^{\alpha}}{x^{(\gamma+1)}}$$

The relative uncertainty on z, assuming uncorrelated errors is:

$$\left(\frac{\sigma_z}{z}\right)^2 = \left(\frac{\sigma_a}{a}\right)^2 + \left(\alpha\frac{\sigma_b}{b}\right)^2 + \left((\gamma+1)\frac{\sigma_x}{x}\right)^2$$

All written directly, with no messy algebraic manipulation required!

But what about a formula like this one?

$$M_{tot} = \frac{P\left(K_1 + K_2\right)^3}{2\pi G \sin^3 i}$$

The sum of the K's cubed and $\sin^3 i$ make propagating the errors ugly. Or do they? Try this simplification of the problem: compute the uncertainty of the *sum of the K's,*

$$\sigma^2_{K_1+K_2} = \sigma^2_{K_1} + \sigma^2_{K_2}$$

and leave *sin(i)* term as is, instead estimating $\sigma_{\sin(i)}$. The relative uncertainty on M_{tot} can now be written as:

$$\left(\frac{\sigma_{M_{tot}}}{M_{tot}}\right)^2 = \left(\frac{\sigma_P}{P}\right)^2 + \left(3\frac{\sigma_{K_1+K_2}}{(K_1+K_2)}\right)^2 + \left(3\frac{\sigma_{\sin i}}{\sin i}\right)^2$$

This is a lot simpler than bashing out the full formula, and it gives the same answer numerically!

The last bit ($\sigma_{\sin i}$) looks like I cheated, but I haven't: the observable quantity really is *sin(i)*, not *i* itself (think about how *i* is measured and why will be come clearer).

One final point: why do astronomers continue to insist on using the archaic and frankly obtuse system of stellar magnitudes instead of physical units like fluxes or luminosities? Among other things, using magnitudes (i.e., logarithmic fluxes or luminosities) makes the error propagation simpler. For example, the uncertainty of a magnitude is just (to within a factor of 1.086) the fractional error in the flux! Thus if you say you want to measure an apparent magnitude to 0.1 mag, you immediately know that you need to measure the fluxes to 0.1% (or 1 part in 1000). Thus magnitude errors are more immediately intuitive than errors on the fluxes.

Similarly, most corrections to measurements of brightness involve attenuation, either geometric attenuation due to the inverse-square law of brightness or line-of-sight absorption (i.e., "optical depth"). Consider the luminosity L_λ you would estimate given an object with an observed flux F_λ a distance D away and behind τ_λ of interstellar extinction:

$$L_\lambda = 4\pi D^2 F_\lambda e^{\tau_\lambda}$$

The observables on the right-hand side all have associated measurement uncertainties, and in this form the uncertainty on your derived L_λ would be something of a mess to be sure, unless you knew the relative uncertainties trick described above.

However, recast as magnitudes,

$$M_\lambda = m_\lambda - 5\log D_{pc} - A_\lambda$$

(note: $A_\lambda = 1.086\tau_\lambda$, and I've converted D into units of parsecs), the uncertainty on the absolute magnitude would be:

$$\sigma_{M_\lambda}^2 = \sigma_{m_\lambda}^2 + \sigma_{A_\lambda}^2 + \left(\frac{5}{\ln 10}\frac{\sigma_D}{D}\right)^2$$

All the uncertainties propagate as the sums of the squares of the various observables (although the last looks a little ugly because of the log10 business).

Rick Pogge (2004 Sept 21)