Glen Cowan
7 February 2006

Numerical error propagation

This note summarizes how to propagate the errors from a fit (e.g., using Least Squares) into a quantity that is only determined numerically from the parameters of the fit.

Suppose the data sample consists of a set of values $\mathbf{x} = (x_1, \ldots, x_N)$, where x is supposed to follow a pdf $f(x;\boldsymbol{\theta})$. Here $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ is the set of parameters to be estimated, and the fitted values (the estimates) are written with hats: $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_n)$. The least-squares or maximum-likelihood fit will also result in an estimate of the covariance matrix for the estimators,

$$V_{ij} = \mathrm{cov}[\hat{\theta}_i, \hat{\theta}_j] \ . \tag{1}$$

Now suppose we are interested in a set of m functions $\boldsymbol{\eta}(\boldsymbol{\theta}) = (\eta_1(\boldsymbol{\theta}), \ldots, \eta_m(\boldsymbol{\theta}))$ which may be calculable only numerically. Very often we may have only a single function ($m = 1$) but in general the value of m is arbitrary. We can estimate $\boldsymbol{\eta}(\boldsymbol{\theta})$ by evaluating the function with the estimate $\hat{\boldsymbol{\theta}}$, i.e., we use

$$\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}(\hat{\boldsymbol{\theta}}) \tag{2}$$

Our goal is to find the covariance matrix

$$U_{ij} = \mathrm{cov}[\hat{\eta}_i, \hat{\eta}_j] \tag{3}$$

as a function of the original covariances V_{ij}, i.e., we want to propagate the errors in $\hat{\boldsymbol{\theta}}$ into those of $\hat{\boldsymbol{\eta}}$. As long as the functions $\boldsymbol{\eta}(\boldsymbol{\theta})$ are sufficiently linear in a region of $\boldsymbol{\theta}$-space comparable to the standard deviations of the estimators $\hat{\theta}_i$, then linear error propagation can be used (see, e.g., [1, 2]). This prescription tells us to take

$$U_{ij} = \sum_{k,l=1}^{n} \frac{\partial \hat{\eta}_i}{\partial \hat{\theta}_k} \frac{\partial \hat{\eta}_j}{\partial \hat{\theta}_l} V_{kl} \ . \tag{4}$$

Now the only difficulty is in finding the derivatives appearing in (3) when the functions $\boldsymbol{\eta}(\boldsymbol{\theta})$ can only be evaluated numerically. The solution is clearly to estimate the derivatives using a simple finite difference technique, i.e., we take

$$\frac{\partial \hat{\eta}_i}{\partial \hat{\theta}_k} \approx \frac{\eta_i(\hat{\theta}_1, \ldots, \hat{\theta}_k + \Delta\theta_k, \ldots, \hat{\theta}_n) - \eta_i(\hat{\theta}_1, \ldots, \hat{\theta}_k - \Delta\theta_k, \ldots, \hat{\theta}_n)}{2\Delta\theta_k} \tag{5}$$

for some appropriately chosen step size $\Delta\theta_k$. This step should not be so small that one is sensitive to round-off or other numerical errors, and not too large that the any nonlinearities in the function are important. One could take, e.g.,

$$\Delta\theta_k = c\sqrt{V_{kk}} \ , \tag{6}$$

where c is a constant not greater than unity (perhaps 0.1 to 0.5). If the nonlinear nature of the function is such that this would result in a poor estimate of the derivative, then the entire procedure of linear error propagation is invalid.

References

[1] Eidelman et al., Physics Letters B592, 1 (2004), Section 32.1.4; also available from `pdg.lbl.gov`.

[2] G. Cowan, *Statistical Data Analysis*, Oxford University Press, 1998, Section 1.6.