The Jackknife Approach

Jackknife estimators allow to correct for a bias and its statistical error. The method was introduced in the 1950s in papers by Quenouille and Tukey. The jackknife method is **recommended as the standard** for error bar calculations. In unbiased situations the jackknife and the usual error bars agree. Otherwise the jackknife estimates are improvements, so that one cannot loose. In particular, the jackknife method solves the question of error propagation elegantly and with little efforts involved.

The unbiased estimator of the expectation value $\widehat{x}$ is

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Normally bias problems occur when one estimates a non-linear function of $\widehat{x}$:

$$\widehat{f} = f(\widehat{x}) \ . \tag{1}$$

Typically, the bias is of order $1/N$:

$$\text{bias}\ (\overline{f}) \ = \ \widehat{f} - \langle \overline{f} \rangle \ = \ \frac{a_1}{N} + \frac{a_2}{N^2} + O(\frac{1}{N^3})\ . \tag{2}$$

Unfortunately, we lost the ability to estimate the variance $\sigma^2(\overline{f}) = \sigma^2(f)/N$ via the standard equation

$$s^2(\overline{f}) \ = \ \frac{1}{N} s^2(f) \ = \ \frac{1}{N\,(N-1)} \sum_{i=1}^{N} (f_i - \overline{f})^2\ , \tag{3}$$

because $f_i = f(x_i)$ is not a valid estimator of $\widehat{f}$: $\widehat{f} - \langle f_i \rangle = \mathcal{O}(1)$. Also it is in non-trivial applications almost always a bad idea to to use standard error propagation formulas with the aim to deduce $\triangle \overline{f}$ from $\triangle \overline{x}$. Jackknife methods are not only easier to implement, but also more precise and far more **robust**.

The error bar problem for the estimator $\overline{f}$ is conveniently overcome by using

jackknife estimators $\overline{f}^J$, f_i^J, defined by

$$\overline{f}^J = \frac{1}{N}\sum_{i=1}^{N} f_i^J \text{ with } f_i^J = f(x_i^J) \text{ and } x_i^J = \frac{1}{N-1}\sum_{k\neq i} x_k \,. \tag{4}$$

The estimator for the variance $\sigma^2(\overline{f}^J)$ is

$$s_J^2(\overline{f}^J) = \frac{N-1}{N}\sum_{i=1}^{N}(f_i^J - \overline{f}^J)^2 \,. \tag{5}$$

Straightforward algebra shows that in the unbiased case the estimator of the jackknife variance (5) reduces to the normal variance (3).

Note that only of order N (not N^2) operations are needed to construct the jackknife averages $x_i^J,\ i = 1, \ldots, N$ from the original data.
