Corrected error calculation for iterative Bayesian unfolding

Tim Adye, Rutherford Appleton Laboratory

4th February 2011

The unfolding method based on iterative application of Bayes' theorem described by D'Agostini [1] (though similar to the iterative procedure of Mülthei and Schorr [2]) is a convenient method, popular in Particle Physics.

Measurement uncertainties

As with all unfolding methods, it is important to understand the uncertainties in the unfolded distribution, and especially the bin-to-bin correlations that ensue as a result of the regularisation process (in the Bayes method without additional smoothing, regularisation comes about as a result of limiting the number of iterations). In many cases, the largest source of uncertainty is from propagation of the measurement uncertainties through the unfolding matrix.

D'Agostini ([1] section 4) gives the unfolded distribution ("estimated causes"), $\hat{n}(\mathrm{C}_i)$, as the result of applying the unfolding matrix, M_{ij}, to the measurements ("effects"), $n(\mathrm{E}_j)$:

$$\hat{n}(\mathrm{C}_i) = \sum_{j=1}^{n_\mathrm{E}} M_{ij} n(\mathrm{E}_j) \tag{1}$$

where

$$M_{ij} = \frac{P(\mathrm{E}_j|\mathrm{C}_i) P_0(\mathrm{C}_i)}{[\sum_{l=1}^{n_\mathrm{E}} P(\mathrm{E}_l|\mathrm{C}_i)][\sum_{l=1}^{n_\mathrm{C}} P(\mathrm{E}_j|\mathrm{C}_l) P_0(C_l)]} \tag{2}$$

and $P(\mathrm{E}_j|\mathrm{C}_i)$ is the response matrix, $\epsilon_i \equiv \sum_{j=1}^{n_\mathrm{E}} P(\mathrm{E}_j|\mathrm{C}_i)$ are the efficiencies, and $P_0(C_l)$ is the prior distribution — initially arbitrary (eg. flat or MC model), but updated on subsequent iterations.

D'Agostini then calculates the covariance matrix, which here we call $V(\hat{n}(\mathrm{C}_k), \hat{n}(\mathrm{C}_l))$, by error propagation from $n(\mathrm{E}_j)$, but assumes that M_{ij} is itself independent of $n(\mathrm{E}_j)$. That is only true for the first iteration. For subsequent iterations, $P_0(\mathrm{C}_i)$ is replaced by $\hat{n}(\mathrm{C}_i)/\hat{N}_\mathrm{true}$ ($\hat{N}_\mathrm{true} \equiv \sum_{i=1}^{n_\mathrm{C}} \hat{n}(\mathrm{C}_i)$) from the previous iteration, and $\hat{n}(\mathrm{C}_i)$ depends on $n(\mathrm{E}_j)$ (eq. 1).

To take this into account, we compute the error propagation matrix

$$\frac{\partial \hat{n}(\mathrm{C}_i)}{\partial n(\mathrm{E}_j)} = M_{ij} + \sum_{k=1}^{n_\mathrm{E}} M_{ik} n(\mathrm{E}_k) \left(\frac{1}{n_0(\mathrm{C}_i)} \frac{\partial n_0(\mathrm{C}_i)}{\partial n(\mathrm{E}_j)} - \sum_{l=1}^{n_\mathrm{C}} \frac{\epsilon_l}{n_0(\mathrm{C}_l)} \frac{\partial n_0(\mathrm{C}_l)}{\partial n(\mathrm{E}_j)} M_{lk} \right) \tag{3}$$

This depends upon the matrix $\frac{\partial n_0(\mathrm{C}_i)}{\partial n(\mathrm{E}_j)}$ which is $\frac{\partial \hat{n}(\mathrm{C}_i)}{\partial n(\mathrm{E}_j)}$ from the previous iteration. In the first iteration, the second term vanishes ($\frac{\partial n_0(\mathrm{C}_i)}{\partial n(\mathrm{E}_j)} = 0$) and we get $\frac{\partial \hat{n}(\mathrm{C}_i)}{\partial n(\mathrm{E}_j)} = M_{ij}$.

We can use the error propagation matrix to obtain the covariance matrix on the unfolded distribution

$$V(\hat{n}(\mathrm{C}_k), \hat{n}(\mathrm{C}_l)) = \sum_{i,j=1}^{n_\mathrm{E}} \frac{\partial \hat{n}(\mathrm{C}_k)}{\partial n(\mathrm{E}_i)} V(n(\mathrm{E}_i), n(\mathrm{E}_j)) \frac{\partial \hat{n}(\mathrm{C}_l)}{\partial n(\mathrm{E}_j)} \tag{4}$$

from the covariance matrix of the measurements, $V(n(\mathrm{E}_i), n(\mathrm{E}_j))$.

This new formula has been compared to the results of toy MC tests and agrees well. Without the new second term, the error is underestimated if more than one iteration is used — by around 20% per iteration in some cases.

D'Agostini takes a multinomial distribution for the bin contents, and hence

$$V(n(\mathrm{E}_i), n(\mathrm{E}_j)) = n(\mathrm{E}_i)\delta_{ij} - \frac{n(\mathrm{E}_i)n(\mathrm{E}_j)}{\hat{N}_\mathrm{true}} \tag{5}$$

That describes a histogram with the fixed normalisation, ie. fixed total number of measured events. On the other hand, in counting experiments common in particle physics, each bin is independently Poisson distributed, with

$$V(n(\mathrm{E}_i), n(\mathrm{E}_j)) = n(\mathrm{E}_i)\delta_{ij} \tag{6}$$

Other, arbitrary, bin errors (perhaps even correlated) may also be used in equation 4.

Response matrix uncertainties

The response matrix, $P(\mathrm{E}_j|\mathrm{C}_i)$, is usually estimated by Monte Carlo. If only limited MC statistics are available, then there will be uncertainties on these terms. Their effect can be determined using

$$\frac{\partial \hat{n}(\mathrm{C}_i)}{\partial P(\mathrm{E}_j|\mathrm{C}_k)} = \frac{n_0(\mathrm{C}_i)n(\mathrm{E}_j)}{f_j \epsilon_i} \left(\delta_{ik} - \frac{n_0(\mathrm{C}_k)}{f_j} \right) \tag{7}$$

where here $f_j \equiv \sum_{l=1}^{n_\mathrm{C}} P(\mathrm{E}_j|\mathrm{C}_l) n_0(\mathrm{C}_l)$ is the folded prior. The covariance matrix due to these errors is given by

$$V(\hat{n}(\mathrm{C}_k), \hat{n}(\mathrm{C}_l)) = \sum_{i,r=1}^{n_\mathrm{E}} \sum_{j,s=1}^{n_\mathrm{C}} \frac{\partial \hat{n}(\mathrm{C}_k)}{\partial P(\mathrm{E}_i|\mathrm{C}_j)} V(P(\mathrm{E}_i|\mathrm{C}_j), P(\mathrm{E}_r|\mathrm{C}_s)) \frac{\partial \hat{n}(\mathrm{C}_l)}{\partial P(\mathrm{E}_r|\mathrm{C}_s)} \tag{8}$$

where $V(P(\mathrm{E}_i|\mathrm{C}_j), P(\mathrm{E}_r|\mathrm{C}_s))$ can be taken as multinomial, Poisson, or other distribution.

References

[1] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," Nucl. Instrum. Meth. A **362** (1995) 487.

[2] H. N. Mülthei and B. Schorr, "On an Iterative Method for the Unfolding of Spectra," Nucl. Instrum. Meth. A **257** (1987) 371.