Effect of Correlations on Propagated Uncertainties

Frank Porter

January 26, 2006

Consider the measurement of a branching fraction. Perhaps this is obtained via fitting a distribution to a signal piece and a background piece. The basic quantity of interest in the fit is the signal yield, y. However, there are various other parameters, $\beta = \{\beta_i\}$, that may be involved in defining the signal and background shapes. These other parameters are in general not precisely known. They may be estimated, including an error matrix, from some subsidiary measurements.

Sometimes the "shape" parameters are allowed to float in the fit. In this case, the uncertainties (including correlations) in these parameters may be folded into the fit results. But sometimes, the analysis is performed by holding these parameters fixed, at best-estimate values $\hat{\beta}_i$. In that case, the uncertainty induced in y due to the uncertainty in β is incorporated later as a "systematic error".

This systematic uncertainty is sometimes evaluated by selecting each of the β_i in turn. The fit for signal yield is redone with $\beta_i = \hat{\beta}_i \pm \sigma_{\beta_i}$. The total systematic is evaluated by adding all of the resulting variations in y in quadrature.

The purpose of this note is to discuss the issue of correlations among the estimators for β, which are neglected in this procedure. We have $y = y(\beta)$.

We'll work in the context of simple linear propagation of errors. In general the variance on y, due to fluctuations in β, is estimated by:

$$(\delta y)^2 \approx \sum_i \sum_j (M_\alpha)_{ij} \frac{\partial y}{\partial \beta_i}\Big|_{\beta=\hat{\beta}} \frac{\partial y}{\partial \beta_j}\Big|_{\beta=\hat{\beta}},$$

where M is the error matrix. When we fit the signal yield with all of the other parameters fixed, the fit does not give any of the terms in this sum. The systematic uncertainty procedure described above estimates the $i = j$ terms. However, it omits the $i \neq j$ terms, that is the correlations among the estimates β are neglected.

In many situations, this neglect is justified. On the other hand, it isn't justified in principle, and there can be situations where it makes a difference.

For example, consider the case with two β parameters, with standard deviations σ_1 and σ_2, and correlation coefficient ρ. The neglected contribution is:

$$2\rho\sigma_1\sigma_2 \frac{\partial y}{\partial \beta_1}\Big|_{\beta=\hat{\beta}} \frac{\partial y}{\partial \beta_2}\Big|_{\beta=\hat{\beta}}.$$

To evaluate how important this is, we compare with

$$\sigma_1^2 \left[\frac{\partial y}{\partial \beta_1}\Big|_{\beta=\hat{\beta}} \right]^2 + \sigma_2^2 \left[\frac{\partial y}{\partial \beta_2}\Big|_{\beta=\hat{\beta}} \right]^2.$$

To make the illustration simple, suppose the derivatives with respect to β_1 and β_2 are equal. Then, we may compare

$$\sigma_1^2 + \sigma_2^2$$

with the result which also includes the $\beta_1 - \beta_2$ correlation:

$$\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2.$$

Since $-1 \leq \rho \leq +1$, this latter result lies in the range

$$(\sigma_1 - \sigma_2)^2 \leq \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \leq (\sigma_1 + \sigma_2)^2.$$

The worst case, with $\sigma_1 = \sigma_2 = \sigma$ is an error of 2σ, to be compared with $\sqrt{2}\sigma$. The systematic uncertainty may be underestimated by as much as 40% if correlations among the two β parameters are neglected (it could also be overestimated, by as much as 100%, depending on ρ).

For > 2 parameters β, the error can be even greater, if the correlations choose to conspire against you.