Tag & Probe parametrizations

m.mieskolainen@imperial.ac.uk, October 29, 2024

We are not being overly pedantic here. Technical code details, such as the PDF integral normalization to unity within the fit range and the histogram binwidth normalization (discrete measure), are implicit in the notation.

Code: github.com/mieskolainen/icenet/tree/master/icefit

1 Parametrization I: $(\theta_S^{\mathbf{pass}}, \theta_B^{\mathbf{pass}}) \otimes (\theta_S^{\mathbf{fail}}, \theta_B^{\mathbf{fail}})$, two independent histogram fits

Here we use two independent equations, for passing and failing event selection histograms

$$\begin{pmatrix} C_{\text{tot}}^{\text{pass}} \\ C_{\text{tot}}^{\text{fail}} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \theta_S^{\text{pass}} \\ \theta_B^{\text{pass}} \\ \theta_S^{\text{fail}} \\ \theta_B^{\text{fail}} \end{pmatrix}, \tag{1}$$

where $\theta_i^j \geq 0$ represents signal and background yield parameters (4) and the observed event counts C_{tot}^j obey

$$C_{\text{tot}} \equiv C_{\text{tot}}^{\text{pass}} + C_{\text{tot}}^{\text{fail}}. \tag{2}$$

After executing the fit, the signal efficiency estimate and its error propagated uncertainty are

$$\epsilon_S = \theta_S^{\text{pass}} / (\theta_S^{\text{pass}} + \theta_S^{\text{fail}}) \tag{3}$$

$$\sigma^2(\epsilon_S) = d_x^2 \sigma^2(\theta_S^{\text{pass}}) + d_y^2 \sigma^2(\theta_S^{\text{fail}}), \tag{4}$$

where derivative variables are

$$d_x = \theta_S^{\text{fail}} / (\theta_S^{\text{pass}} + \theta_S^{\text{fail}})^2 \tag{5}$$

$$d_y = -\theta_S^{\text{pass}} / (\theta_S^{\text{pass}} + \theta_S^{\text{fail}})^2. \tag{6}$$

This can be derived using error propagation on a ratio $f \equiv x/(x+y)$. No correlation term is needed because fits were independent and the uncertainties squared $\sigma^2(\theta_S^{\text{pass}})$ and $\sigma^2(\theta_S^{\text{fail}})$ are from the fitting routine.

2 Parametrization II: $(\epsilon_S, \epsilon_B, \theta_S, \theta_B)$, one simultaneous fit of two histograms

We start by re-parametrizing with efficiencies ϵ_i by coupling the equations, and obtain a system of two equations

$$\begin{pmatrix} C_{\mathrm{tot}}^{\mathrm{pass}} \\ C_{\mathrm{tot}}^{\mathrm{fail}} \end{pmatrix} = \begin{pmatrix} \epsilon_S & \epsilon_B \\ 1-\epsilon_S & 1-\epsilon_B \end{pmatrix} \begin{pmatrix} \theta_S \\ \theta_B \end{pmatrix}. \tag{7}$$

The determinant of the matrix is $\epsilon_S - \epsilon_B$. Thus, the equation is formally well defined when $\epsilon_S \neq \epsilon_B$. The efficiency parameter bounds are $\epsilon_S, \epsilon_B \in [0,1]$.

We could simply invert the matrix, if the efficiencies would be known. In general, we need shape information and non-linear fitting to solve all 4 free parameters. The signal efficiency ϵ_S and its (expanded) uncertainty are provided now directly by the non-linear fitting routine (e.g. Minuit).

3 'Unitary' parametrization I: $(\epsilon_S, \epsilon_B, \theta_S)$, one simultaneous fit of two histograms with one constraint

If we use the constraint

$$C_{\mathrm{tot}} = \theta_S + \theta_B, \tag{8}$$

i.e. yield extractions for signal and background must sum to the total observed counts, then we can reduce one free parameter by rewriting the system as

$$\begin{pmatrix} C_{\mathrm{tot}}^{\mathrm{pass}} \\ C_{\mathrm{tot}}^{\mathrm{fail}} \end{pmatrix} = \begin{pmatrix} \epsilon_S & \epsilon_B \\ 1-\epsilon_S & 1-\epsilon_B \end{pmatrix} \begin{pmatrix} \theta_S \\ C_{\mathrm{tot}} - \theta_S \end{pmatrix}. \tag{9}$$

Obtaining the 3 free parameters and their uncertainties is done using the non-linear fit.

4 'Unitary' parametrization II: (ϵ_S, f_S), one simultaneous fit of two histograms with two constraints

This is an extension of the previous one, by using a total signal fraction $f_S \in [0,1]$ parameter, which gives

$$C_{\mathrm{tot}} = \theta_S + \theta_B = f_S C_{\mathrm{tot}} + (1-f_S) C_{\mathrm{tot}}. \tag{10}$$

Now, we can eliminate ϵ_B by solving it as

$$\epsilon_B = \frac{\Omega - \epsilon_S f_S}{1 - f_S}, \tag{11}$$

where we defined the variable $\Omega \equiv \frac{C_{\mathrm{tot}}^{\mathrm{pass}}}{C_{\mathrm{tot}}}$. The system is finally

$$\begin{pmatrix} C_{\mathrm{tot}}^{\mathrm{pass}} \\ C_{\mathrm{tot}}^{\mathrm{fail}} \end{pmatrix} = \begin{pmatrix} \epsilon_S & \Omega - \epsilon_S f_S \\ 1-\epsilon_S & 1 - f_S - \Omega + \epsilon_S f_S \end{pmatrix} \begin{pmatrix} f_S C_{\mathrm{tot}} \\ C_{\mathrm{tot}} \end{pmatrix}. \tag{12}$$

With only 2 free parameters (ϵ_S, f_S) to be extracted by the non-linear fit. The determinant of the matrix is $\epsilon_S - \Omega$. In practise when fitting, one may need to bound the matrix elements to be non-negative, e.g., $\Omega - \epsilon_S f_S \to \max(0, \Omega - \epsilon_S f_S)$.