Glen Cowan
RHUL Physics
7 April 2018

Approximate error on peak position for low s/b

This note gives an approximate formula for the statistical error on the position of a Gaussian peak of width σ on a flat background. The approximation holds in the limit where the expected number of signal events s is large compared to unity but small compared to the expected number of background events under the peak, which can be taken as the number contained in one σ.

Let x be the variable measured for each event and whose spectrum has a peak of width σ and position μ, e.g., here x represents the invariant mass of a photon pair from a Higgs decay. Suppose $-a \leq x \leq a$ with $a \gg \mu \gg \sigma$, i.e., the entire peak is well contained within the flat background. Under these conditions the final result will be independent of a.

The probability density function of x can be written as a mixture of uniform and Gaussian pdfs as

$$f(x;\mu) = \frac{b}{s+b}\frac{1}{2a} + \frac{s}{s+b}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2} \;, \tag{1}$$

where s is the total expected number of signal events and b is the expected number of background events in $[-a,a]$. One could model the observed number of events as Poisson distributed with a mean of $s+b$, but for purposes of this calculation it is good enough to take $n = s+b$.

For a sample of n independent events the likelihood function is

$$L(\mu) = \prod_{i=1}^{n} \left[\frac{b}{s+b}\frac{1}{2a} + \frac{s}{s+b}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x_i-\mu)^2/2\sigma^2} \right] \;. \tag{2}$$

The log-likelihood is therefore

$$\ln L(\mu) = \sum_{i=1}^{n} \ln\left[1 + \frac{2as}{b}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x_i-\mu)^2/2\sigma^2} \right] + C \;, \tag{3}$$

where C represents terms that do not depend on the parameter μ. The maximum-likelihood estimator $\hat{\mu}$ is found by setting the derivative of $\ln L$ to zero, which in this case must be done numerically.

For a sufficiently large data sample the variance of $\hat{\mu}$ can be found from the second derivative of $\ln L$. For this we can obtain a simple expression in the limit where a small signal peak is sitting on top of a large background. The expected number of background events b is proportional to the (arbitrary) size of the interval, $2a$. We therefore define β as the expected number of background events in one unit of mass resolution σ,

$$\beta = \frac{b\sigma}{2a} \;. \tag{4}$$

Having a small signal on a large background thus means $s \ll \beta$. In this case we can expand the logarithms in $\ln L$ to first-order, which gives

$$\ln L(\mu) \approx \sum_{i=1}^{n} \frac{s}{\beta} \frac{1}{\sqrt{2\pi}} e^{-(x_i-\mu)^2/2\sigma^2} + C \ , \tag{5}$$

and the second derivative is

$$\frac{\partial^2 \ln L}{\partial \mu^2} = -\frac{s}{\sqrt{2\pi}\beta\sigma^2} \sum_{i=1}^{n} \left[1 - \frac{(x_i-\mu)^2}{\sigma^2} \right] e^{-(x_i-\mu)^2/2\sigma^2} \ . \tag{6}$$

The inverse variance of $\hat{\mu}$ is given by the negative expectation value of the second derivative of $\ln L$,

$$\frac{1}{\sigma_{\hat{\mu}}^2} = -E\left[\frac{\partial^2 \ln L}{\partial \mu^2} \right] = \frac{ns}{\sqrt{2\pi}\beta\sigma^2} E\left[\left(1 - \frac{(x-\mu)^2}{\sigma^2} \right) e^{-(x-\mu)^2/2\sigma^2} \right] \ . \tag{7}$$

The expectation value is taken with respect to the pdf from Eq. (1). The terms related to the uniform component of the pdf cancel; for the Gaussian part we can take the limits of integration to $\pm\infty$, i.e, we assume the Gaussian peak is well contained within the interval $-a \leq x \leq a$. The $E[\cdot]$ term then evaluates to $s/[(s+b)2\sqrt{2}]$, and furthermore we can set $n = s + b$. Solving Eq. (7) for the standard deviation $\sigma_{\hat{\mu}}$ gives the final result

$$\sigma_{\hat{\mu}} = 2\pi^{1/4} \frac{\sqrt{\beta}}{s} \sigma \ . \tag{8}$$

This result requires $s \ll \beta$ but one must also have s sufficiently large for the asymptotic approximation for the variance to be valid. The result should be sufficiently accurate for use in optimizing an analysis, but for a final determination of the statistical error on $\hat{\mu}$ one would use the full likelihood to find $\sigma_{\hat{\mu}}$ and/or a confidence interval for μ.