What is the Effective Signal and what is it good for?

Thomas Ullrich, BNL

December 4, 2009

Abstract

The effective signal, S_{eff}, is a useful measure to judge the quality of a signal in the presence of background. It takes into account both, the signal strength and the background level and greatly simplifies the calculation of the significance of the measurement. It is often also called the "background-free equivalent". In the following I derive S_{eff} and discuss the advantages in using it.

1 Introduction

The effective signal, S_{eff}, was originally used in the context of evaluating the signal strength (significance) of resonances but can be applied to any situation where a given signal has to be judged in the presence of an underlying background. Should there be no background in the measurement, then S_{eff} is nothing but the signal, $S = S_{\text{eff}}$.

Let us assume in the following that we conduct some measurement and observe a total count T, which is the sum of the actual signal, S, and a background contribution, B. The actual signal can thus be derived using:

$$S = T - B \tag{1}$$

where B has to evaluated by some means. In the case of quarkonia measurements for example, T would be number of unlike-sign pairs in some mass range, and B can be estimates using the number of like-sign pairs in the same mass range.

The significance of the signal is commonly expressed in terms of signal yield divided by the statistical error of the signal, $S/\delta S$. A signal of $S = 10 \pm 5$ is often referred to as a 2σ signal. When expressing the significance in terms of σ one assumes, not always correctly, that the errors are Gaussian distributed. Using Gaussian error propagation the error of S in Eq. 1 is:

$$\delta S = \sqrt{\left(\frac{\partial S}{\partial T}\delta T\right)^2 + \left(\frac{\partial S}{\partial B}\delta B\right)^2} \tag{2}$$

$$= \sqrt{(\delta T)^2 + (\delta B)^2} \tag{3}$$

Since $T = S + B$ and assuming simple counting statistics, i.e., $\delta S = \sqrt{S}$ and $\delta B = \sqrt{B}$, we obtain the well known expression:

$$\delta S = \sqrt{S + 2B}. \tag{4}$$

Note that in some cases where the background is determined for example from fits or event mixing, the statistical error on the background is often neglected and Eq. 4 becomes $\delta S = \sqrt{S + B}$.

Using Eq. 4, the significance of the measurement becomes

$$\frac{S}{\delta S} = \frac{S}{\sqrt{S + 2B}}. \tag{5}$$

2 Background free equivalent

Let us assume a typical measurement with a given S, T, and B. One can now ask the simple question: what signal in the absence of any background yields the same significance. We call this signal the "effective" signal, S_{eff}, or the background-free equivalent. Without background the significance of S_{eff} is simply given by:

$$\frac{S_{\text{eff}}}{\delta S_{\text{eff}}} = \frac{S_{\text{eff}}}{\sqrt{S_{\text{eff}}}} = \sqrt{S_{\text{eff}}}. \tag{6}$$

Since we require that S_{eff} has the same significance as in the case *with* background

$$\frac{S}{\delta S} = \frac{S_{\text{eff}}}{\delta S_{\text{eff}}} \tag{7}$$

$$\frac{S}{\sqrt{S + 2B}} = \sqrt{S_{\text{eff}}} \tag{8}$$

and we obtain easily an expression for S_{eff}:

$$S_{\text{eff}} = \frac{S}{2\frac{B}{S} + 1}. \tag{9}$$

This is the definition of the effective signal S_{eff}. It is the signal strength in the absence of any background that has the same significance as a measurement *with* background.

So what is it good for? An obvious advantage is that it is trivial to calculate the significance. It is simply $\sqrt{S_{\text{eff}}}$. For example a measurement with $S = 100$ and $B = 100$ ($S/B = 1 : 1$) has $S_{\text{eff}} = 50$. The significance of the signal is $\sqrt{50}\sigma \approx 7\sigma$. Quoting the S_{eff} avoids asking what the referring signal-to-background ratio (S/B) is; no need to provide any specifics on background levels.

The true advantage, however, is in the optimization of cuts or the evaluation or comparison of different methods. Optimizing cuts to improve the signal, S, often causes the background level to increase, resulting in no net gain, or even deteriorating the overall significance. Using and tuning techniques to reduce the background most likely also affects the signal. Since S_{eff} reflects both, signal and background, it is the simplest measure to judge the result of any optimization study. An increase in S_{eff} is a gain, a decrease means you made things worse. It is also useful in predictions. Saying that we will measure $S = N$ rare particles of some type in future runs ignores the backgrounds levels and does not allow the reader to judge the quality of the measurement. Quoting S_{eff} avoids this problem. Should the background be derived by means with no significant statistical errors (see above) S_{eff} becomes $S_{\text{eff}} = S/(B/S+1)$. Note, that S_{eff} of course does not reflect any systematical errors that will further effect the true significance.