Extra Problem 1: Suppose we count a number of events n which is assumed to follow a Poisson distribution with mean $s+b$. Here, suppose the expected number of background events b is known, and we want to estimate the number of signal events s.

(a) Show that the ML estimator for s is

$$\hat{s} = n - b$$

for $n > b$ and is zero otherwise.

(b) Suppose we take $\hat{s} = n - b$ as our estimator, regardless of whether n is greater or less than b. Show that the relative statistical error in $\hat{s}$ is

$$\frac{\sigma_{\hat{s}}}{s} = \frac{\sqrt{s+b}}{s} \, .$$

When we design the event selection criteria for the analysis we determine the signal and background efficiencies, and therefore we fix s and b. Thus in order to minimize the relative uncertain in our measurement of s we should set the selection criteria to maximize $s/\sqrt{s+b}$.

(c) Show that

$$\frac{s}{\sqrt{s+b}} \propto \sqrt{\varepsilon_{\rm s} p_{\rm s}} \, ,$$

where $\varepsilon_{\rm s}$ is the signal efficiency and $p_{\rm s}$ is the signal purity, and find the constant of proportionality. (Recall that the expected number of signal events s is related to the efficiency ε_s through the relation $s = \sigma_s L \varepsilon_s$, where σ_s is the signal cross section and L is the integrated luminosity. A similar relation holds for the background.)

Note that if the goal of the analysis is not to measure s, but rather to discover whether the signal process exists, then we should not minimize the relative uncertainty $\sigma_{\hat{s}}/s$ but rather maximize the expected discovery significance based on the p-value of the background-only hypothesis. In many cases this is roughly equivalent to maximizing $s/\sqrt{b}$.