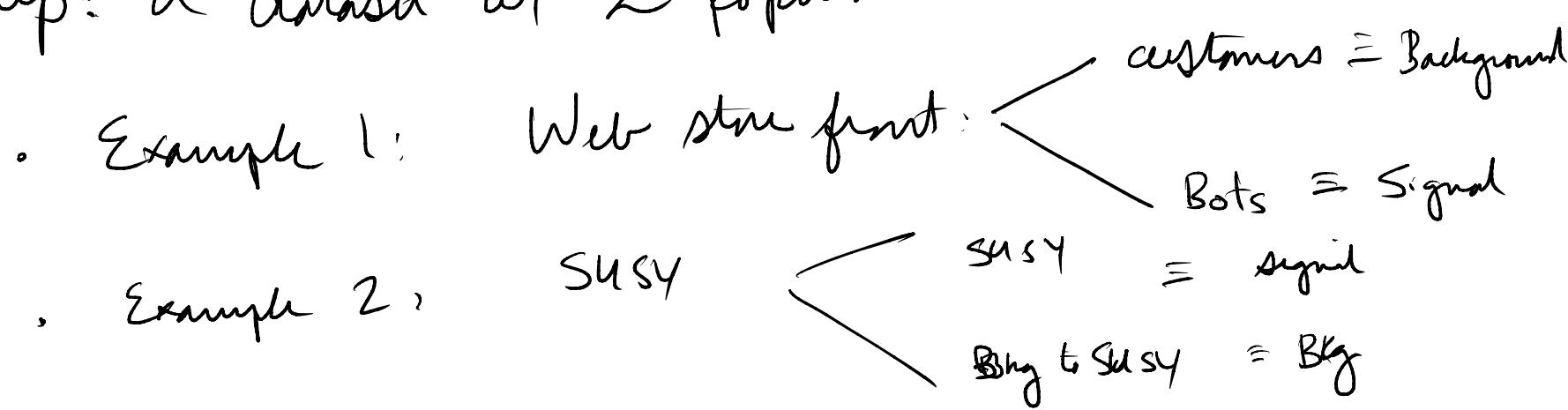


Setup: a dataset w/ 2 populations



⇒ Estimated the expected Background

- Ex 1: previous data
 - Ex 2: Phenix calc + simulation
- $\} \rightarrow \bar{N}_B$

Expect \bar{N}_B

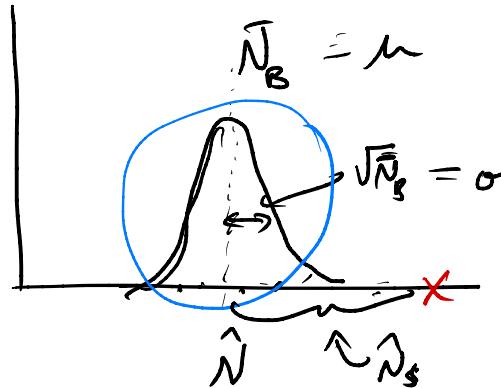
Observe \hat{N}

$$\bar{N}_B = 100 \Rightarrow \sqrt{\bar{N}_B} = 10$$

$\begin{matrix} 90 \\ 100 \\ 110 \end{matrix} \left. \right\} \text{ w/w 1 STD}$

$\begin{matrix} 150 \\ 180 \end{matrix} \rightarrow \text{very unlikely}$

$$\hat{N}_s = 50 \quad \sigma_{\hat{N}_s} = \sqrt{150}$$



Counting \Rightarrow

Poisson Stat

Standard Deviation

if N is the most likely value

\Rightarrow Standard deviation is $\sqrt{N} = \sigma$

$$\begin{aligned} \hat{N}_s &= \hat{N} - \bar{N}_B && \text{very good estimate} \\ \sigma_{\hat{N}_s} &= \text{STD } \hat{N}_s = \sqrt{\hat{N}} && \Rightarrow \text{almost no error} \\ &&& \approx \sqrt{\hat{N}_s + \bar{N}_B} \end{aligned}$$

significance = $\frac{\hat{N}_s}{\sigma_{\hat{N}_s}} \approx \frac{\hat{N}_s}{\sqrt{\hat{N}_s + \bar{N}_B}}$

> 4	observed
> 5	discrepancy

Imagine that there is a observable (feature, raw feature)

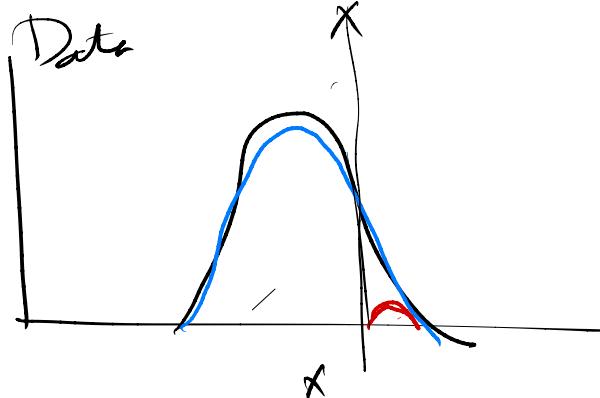
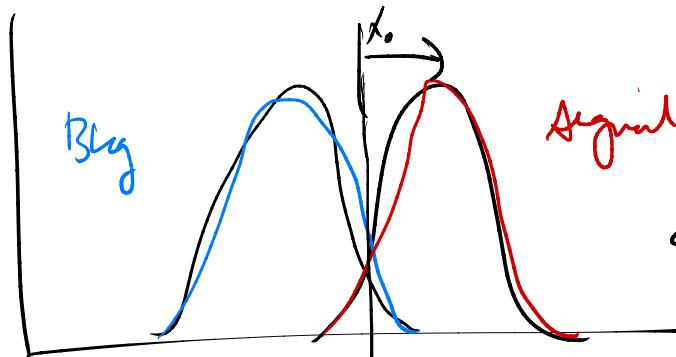
x

True positive rate (TPR)

$E_s = \text{fraction of signal events}$
Passing select of $x > x_0$

False Positive Rate (FPR) in

$E_B = \frac{\text{Bkg}}{\text{Bkg}}$



$E_s \gg E_B$

$\overline{N}_B' \xleftarrow{\text{after selection}}$

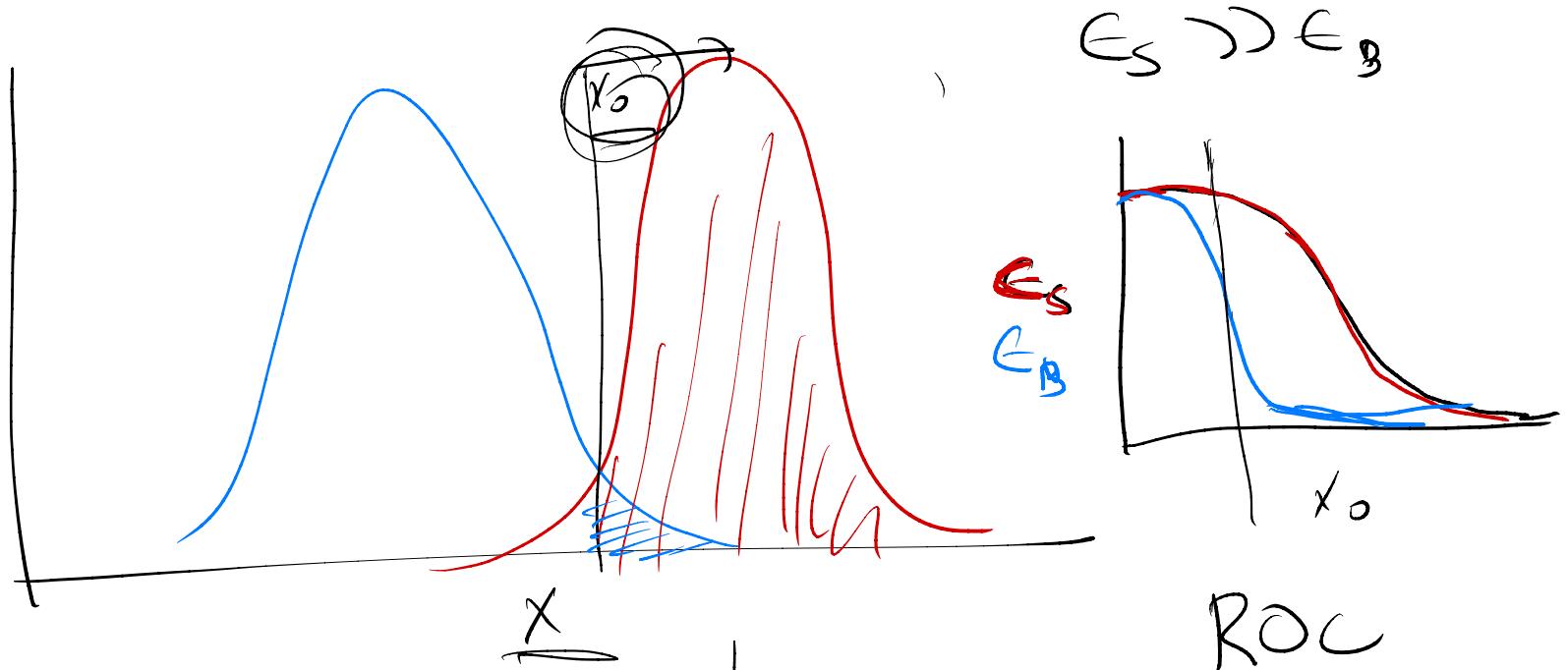
$$\overline{N}_B' = E_B \overline{N}_B \Rightarrow \overline{N}' = \overline{N}_s' + \overline{N}_B' \Leftarrow \text{expected}$$

$$\overline{N}_s' = E_s \overline{N}_s \Rightarrow \hat{N}' \Leftarrow \text{observed}$$

$$\hat{N}_s' = \hat{N}' - \overline{N}_B'$$

$$\sigma_{\hat{N}_s'} = \sqrt{\hat{N}'} \approx \sqrt{\overline{N}'} = \sqrt{E_B \overline{N}_B + E_s \overline{N}_s}$$

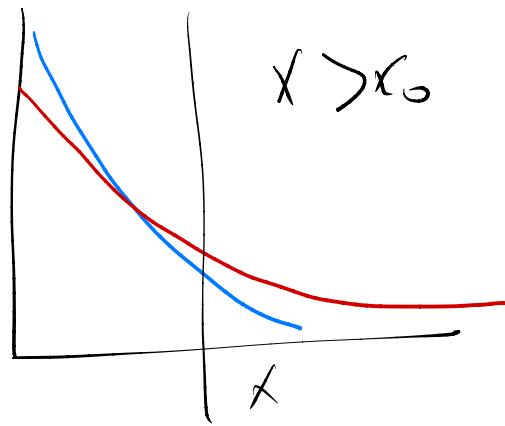
$$\text{significance} = \frac{\bar{N}_S}{\sqrt{\epsilon_S \bar{N}_S + \epsilon_B N_B}} \approx \frac{\epsilon_S \bar{N}_S}{\sqrt{\epsilon_S N_S + \epsilon_B N_B}}$$



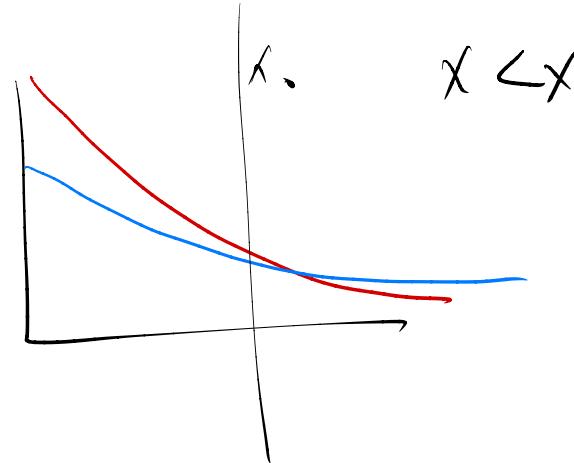
Optimization \Rightarrow select x_0
such that
sig is maximum.



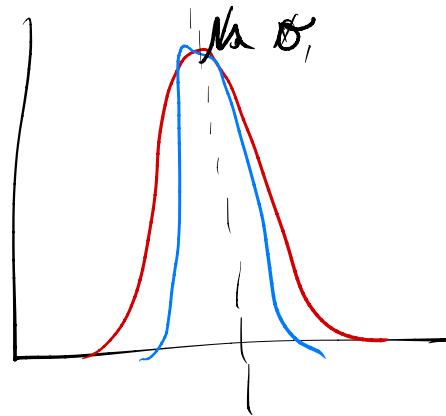
ROC



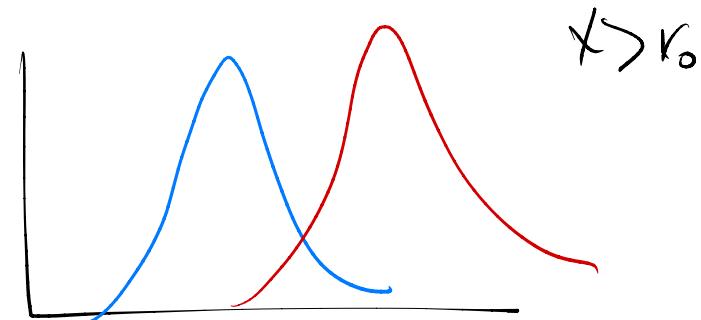
$$x > x_0$$



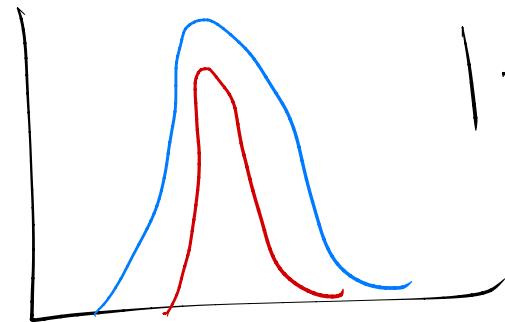
$$x < x_0$$



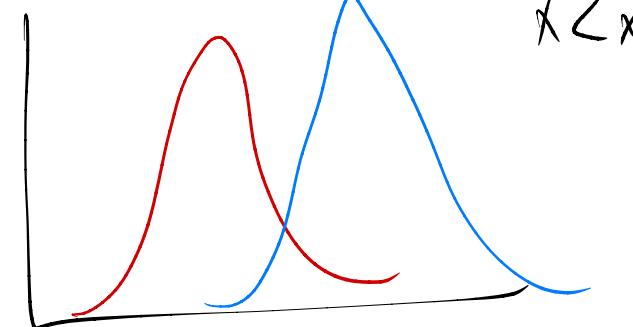
$$|x - \mu| > \sigma$$



$$x > x_0$$



$$|x - \mu| < \sigma$$



$$x < x_0$$