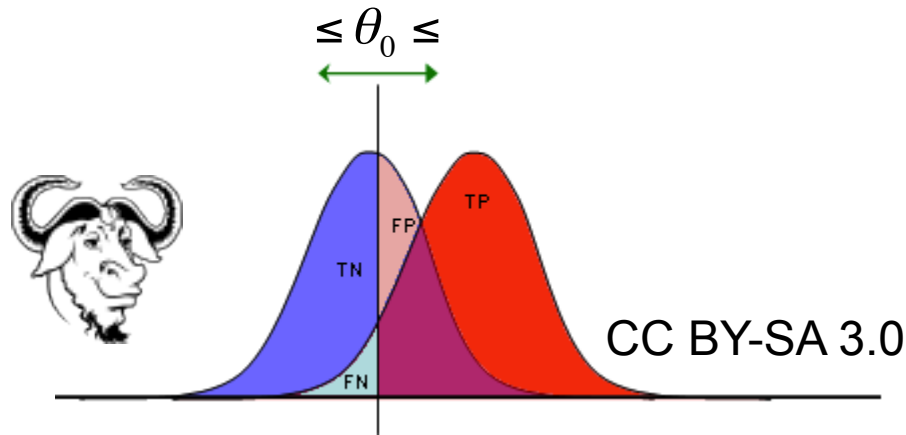


# Continuous Example: Temperature Bound $\Rightarrow$ flue yes / no



Classifier: Decide on disease by varying diagnosis temperature  $\theta_0$ :

What is the “best” classifier value to indicate: you have the flue?

**Red:** true positive (TP)

high temperature and will have flue  
[if a pos. instance is classified as Hit]

**Light blue:** false negative (FN)

too low temp, but will get the flue  
[if a pos. instance is classified as Miss]

**Strong blue:** true negative (TN)

too low temp, and no flue  
[if a neg. instance is classified as Miss]

**Light red:** false positives (FP)

high temperature but never get it  
[if a neg. instance is classified as Hit]

|   |      | Evidence<br>(True Class) |     |
|---|------|--------------------------|-----|
|   |      | pos                      | neg |
| Classifier<br>prediction<br>(Hypothesised<br>class) | Hit  | TP                       | FP  |
|   | Miss | FN                       | TN  |
| Column totals                                       |      | P                        | N   |



# From Confusion table to ROC graph

|                    |   | True class         |                    |
|--------------------|---|--------------------|--------------------|
|                    |   | p                  | n                  |
| Hypothesized class | y | True<br>Positives  | False<br>Positives |
|                    | n | False<br>Negatives | True<br>Negatives  |
| Column totals      |   | P                  | N                  |

$$\text{FP rate} = \frac{\text{FP}}{N}$$

$$\text{TP rate} = \frac{\text{TP}}{P} = \text{Recall}$$

ROC graph uses for x axis == FP rate

ROC graph uses for y axis == TP rate

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{F-score} = \text{Precision} * \text{Recall}$$

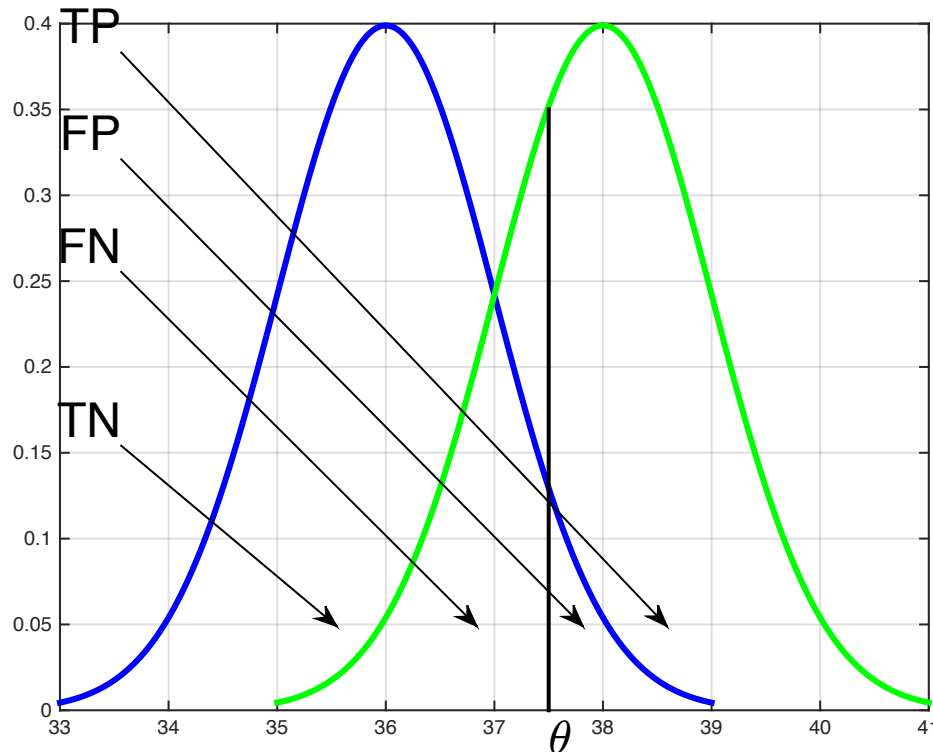
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{P + N}$$

# From Confusion table to ROC graph

|              |                       | Condition<br>(as determined by "Gold standard")   |   |  |
|--------------|-----------------------|---|---|--|
|              |                       | Condition Positive  | Condition Negative  |  |
| Test Outcome | Test Outcome Positive | <b>True Positive</b>  | <b>False Positive</b><br>(Type I error)   | Precision =<br>$\frac{\Sigma \text{ True Positive}}{\Sigma \text{ Test Outcome Positive}}$                 |
|              | Test Outcome Negative | <b>False Negative</b><br>(Type II error)  | <b>True Negative</b>  | Negative predictive value =<br>$\frac{\Sigma \text{ True Negative}}{\Sigma \text{ Test Outcome Negative}}$ |
|              |                       | Sensitivity =<br>$\frac{\Sigma \text{ True Positive}}{\Sigma \text{ Condition Positive}}$ | Specificity =<br>$\frac{\Sigma \text{ True Negative}}{\Sigma \text{ Condition Negative}}$ | Accuracy   |

Co Wikipedia (en):  
„ROC“, CC BY-SA 3.0

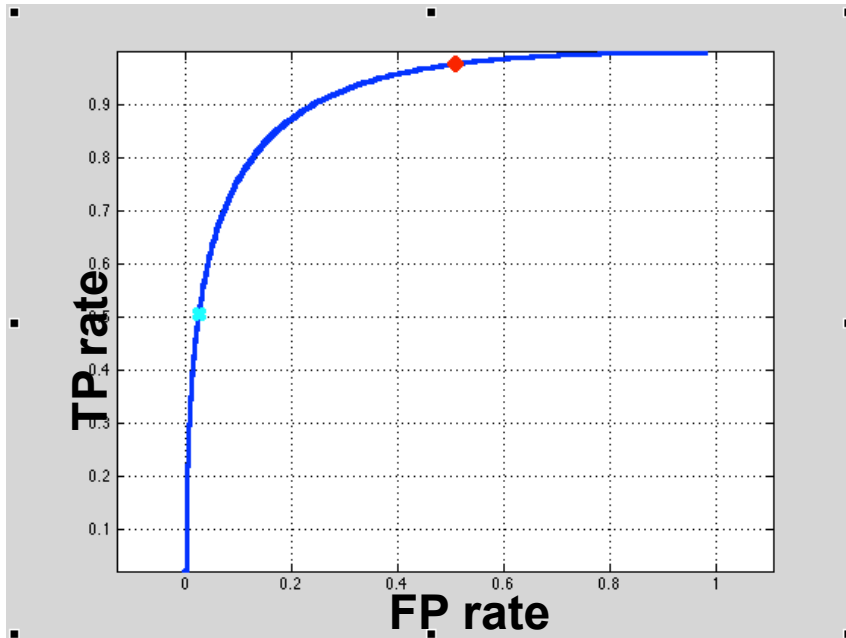
# Data by hand: normpdf



```
%two temperature distributions, healthy and sick
clf;
mu=36;
mu2=38;
sigma=1.0;
sigma2=1.0;
%sample it arround ist modes == peaks
X=linspace(mu-3,mu+3,201);
Y=linspace(mu2-3,mu2+3,201);
plot(X,normpdf(X,mu,sigma),'b','LineWidth',3)
hold on;grid on;
plot(Y,normpdf(Y,mu2,sigma2),'g','LineWidth',3);
plot([37.5 37.5],[0 0.35],'k','LineWidth',2);
% Create textarrow
annotation('textarrow',[0.185
0.376785714285715],...
[0.424167024167026 0.21],'String',
{'TN'},'FontSize',24);
annotation('textarrow',[0.185
0.501785714285714],...
[0.630952380952381 0.21],'String',
{'FN'},'FontSize',24);
annotation('textarrow',[0.185
0.596428571428571],...
[0.764764764764765 0.21],'String',
{'FP'},'FontSize',24);
annotation('textarrow',[0.185
0.671428571428571],...
[0.891891891891892 0.21],'String',
{'TP'},'FontSize',24);
annotation('textbox',...
[0.557707572897 0.0547619047619048
0.0311428571428571 0.05],...
'String','\theta',...
'LineStyle','none',...
'FontSize',24,...
'FontName','Helvetica',...
```

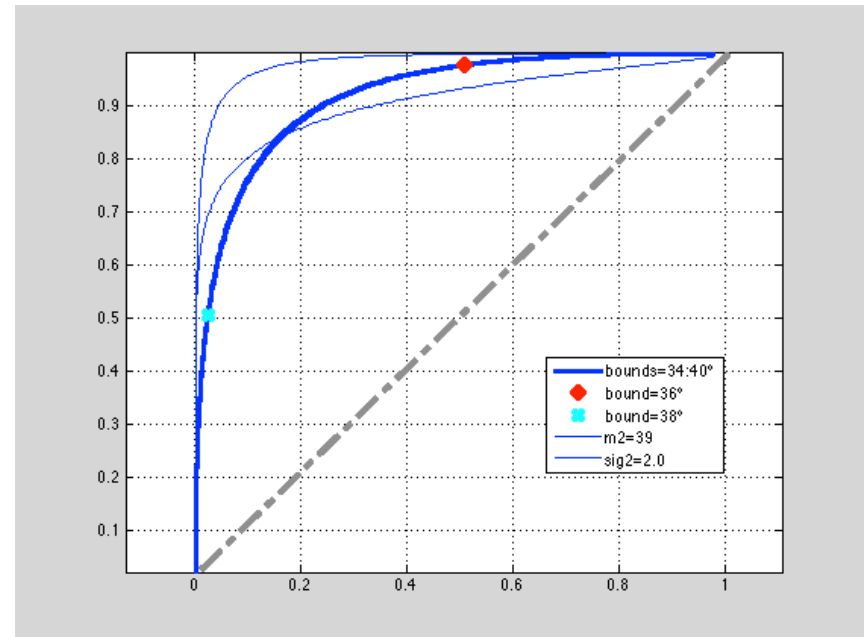
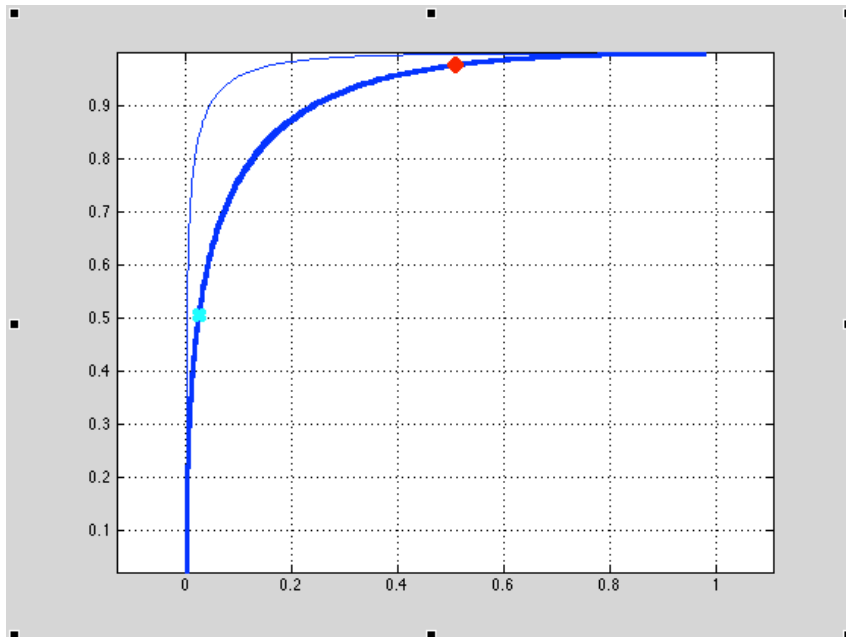


# Results: normcdf, ROC by hand



```
bounds=[34:0.02:40];  
%these are the values for boundaries  
TN=normcdf(bounds,mu,sigma);% BLUE area  
FP=1-TN; % ROSE area  
FN=normcdf(bounds,mu2,sigma2);% CYAN area  
TP=1-FN; % RED area  
TPrate=TP./(TP+FN);% build percentage or rates  
FPrate=FP./(FP+TN);%  
plot(FPrate,TPrate);axis equal;grid on;  
%MY first ROC diagram  
hold on  
plot(FPrate(100),TPrate(100),'rd','LineWidth',3);  
% right in Raw, is left in ROC  
plot(FPrate(200),TPrate(200),'cx','LineWidth',3);  
%
```

# Results: 2nd ROC for $\mu_2=39^\circ$ 3rd ROC for $\sigma_2=2.0$



# Eg.: Application to signals

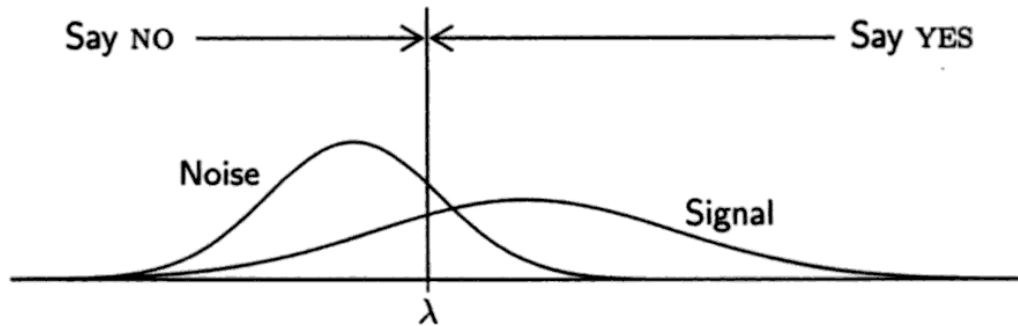


Figure 1.2: The signal and noise distributions of Figure 1.1 shown on a single axis with a decision criterion at the value  $\lambda$ .

$X_n$  : random variable for noise trials, density  $f_n(x)$

$X_s$  : random variable for signal trials, density  $f_s(x)$

false – alarm rate :  $P_F = P(\text{YES} | \text{noise}) =$

$$P(X > \lambda | \text{noise}) = P(X_n > \lambda) = \int_{\lambda}^{\infty} f_n(x) dx$$

$= 1 - F_n(\lambda)$  where  $F_n$  is Cumulative Distribution

Function of  $f_n$