ROC curve

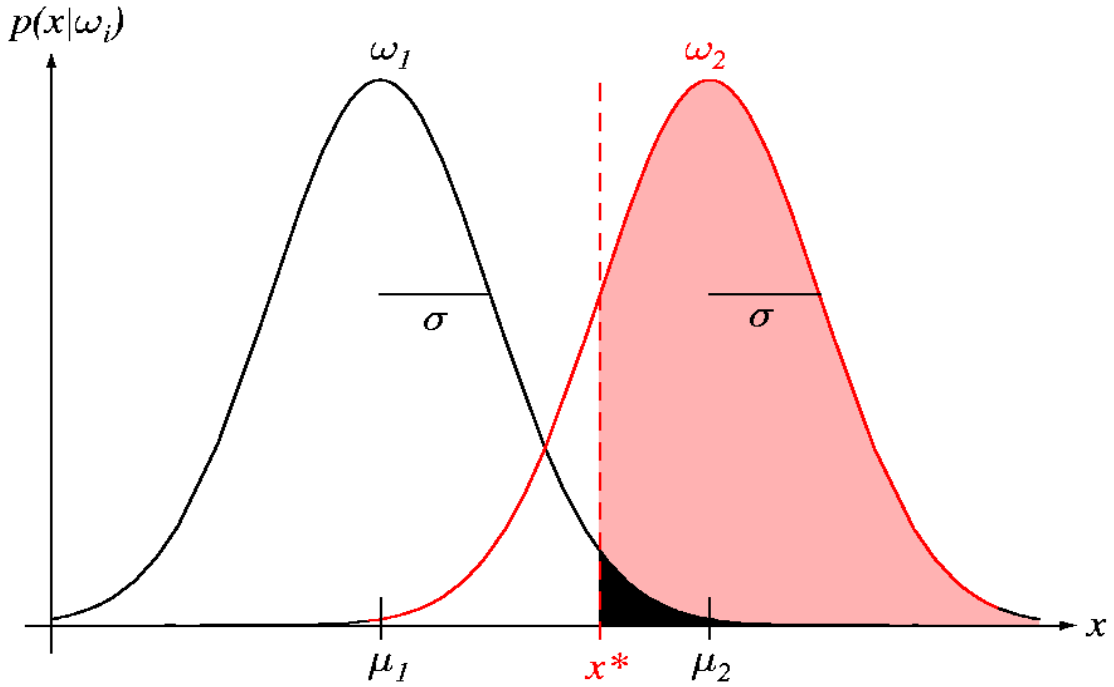
The origin of the abbreviation ROC:

Receiver Operating Characteristic

Classification Error

We consider two normal distributions of equal variance. This situation has applications in psychology and radar detection. Class 1: no object to be detected. Class 2: object is present.

A decision threshold x* will determine the probability of a hit (pink area) and of a false alarm (black area).



from Duda, Hart, Stork (2001) Pattern classification

Classification Error

Four situations may occur:

1) $P(x > x^* | \omega_2)$, *hit*

2) $P(x > x^* | \omega_1)$, *false alarm*

3) $P(x < x^* | \omega_2)$, *miss*

4) $P(x < x^* | \omega_1)$, *correct rejection*

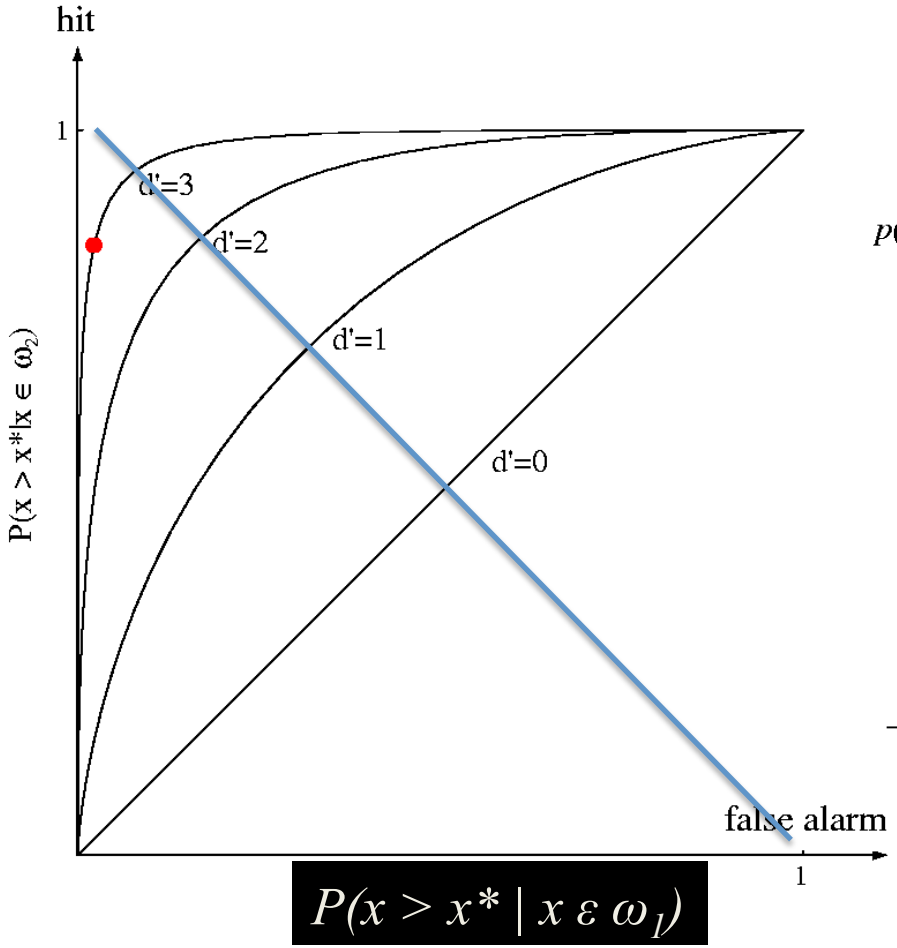
For a large number of trials, the hit and false alarm rates can be determined experimentally.

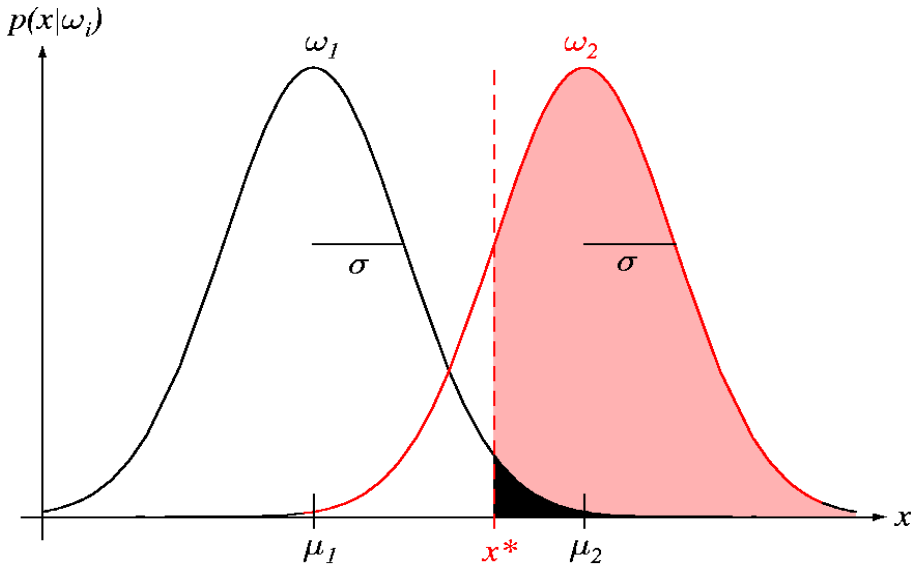
Discriminability

$$d' = \frac{|\mu_2 - \mu_1|}{\sigma}$$

Applies to two normal distributions of equal variance.

Receiver operating characteristic (ROC) curve

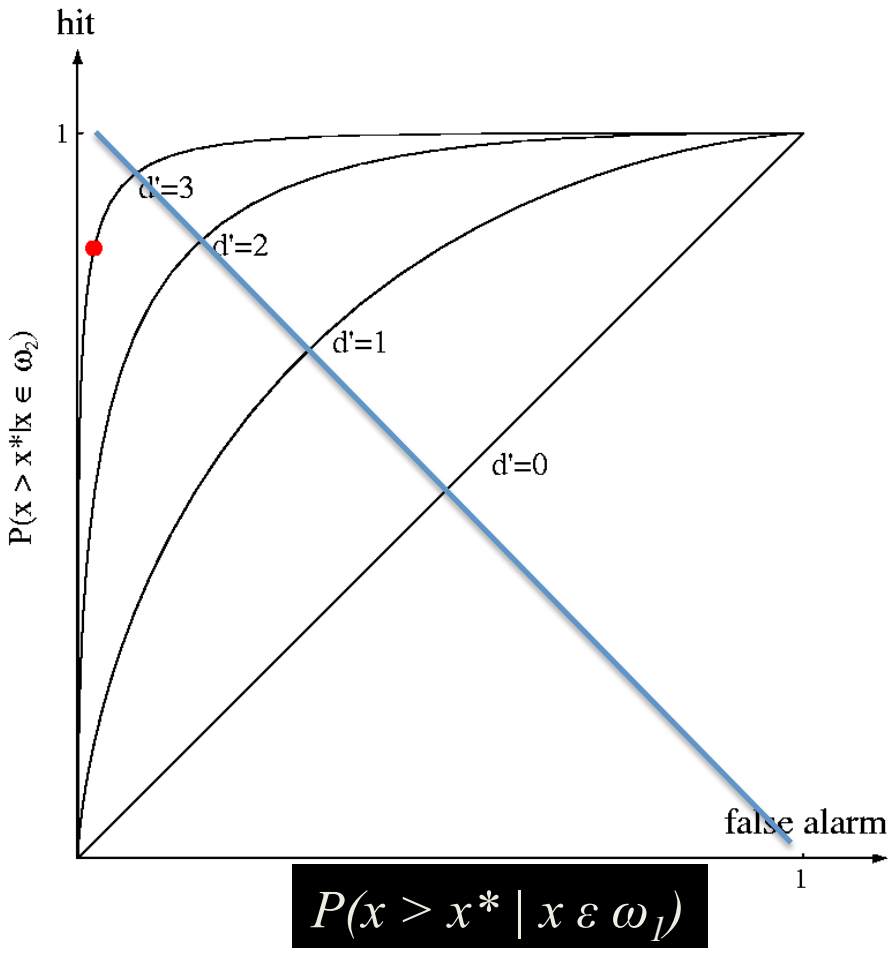




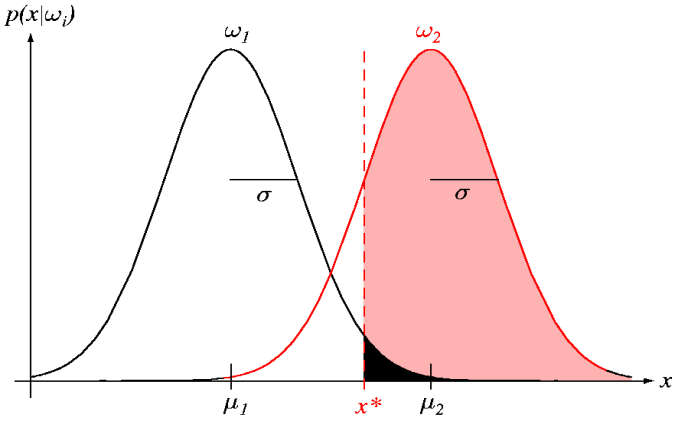
Plot of *hit* vs. *false alarm*

from Duda, Hart, Stork (2001) Pattern classification

Receiver operating characteristic (ROC) curve



Plot of *hit* vs. *false alarm*

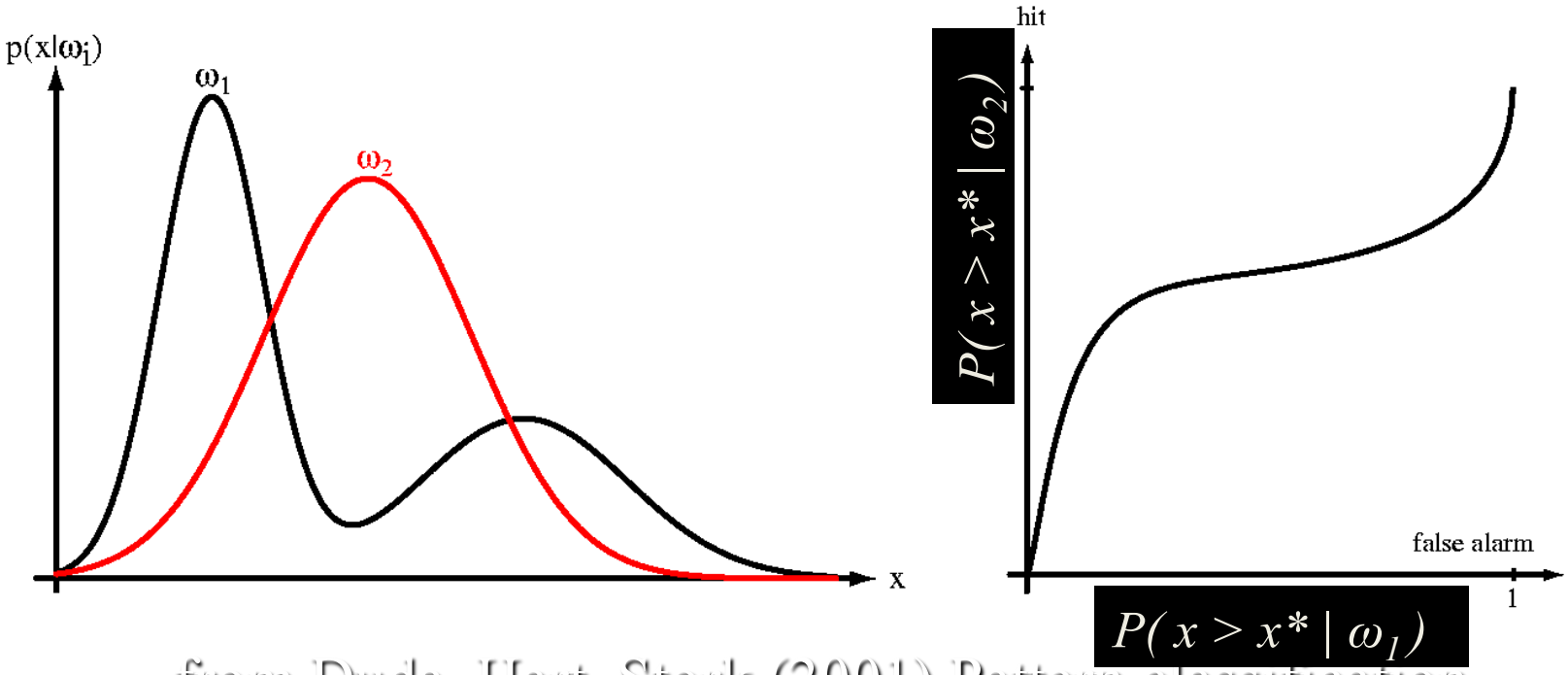


The red dot is a pair of experimentally determined values of false alarm and hit. This point lies on a curve of a given discriminability (here d' = 3).
Note that for an optimal Bayes classifier the values of hit rate and false alarm rate should add up to 1. That point lies at the crossing of the curve and the blue straight line that connects points (1,0) and (0,1).

from Duda, Hart, Stork (2001) Pattern classification

Operating characteristic curve

The OC curves may not be always symmetric. Typically, such a curve is determined experimentally by changing some parameter of a classifier and measuring the hit and false alarm. Such curves are used to compare classifiers.

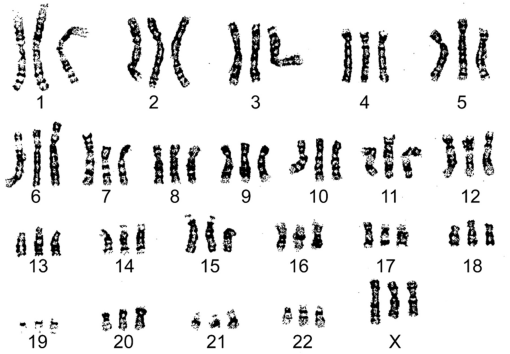


from Duda, Hart, Stork (2001) Pattern classification

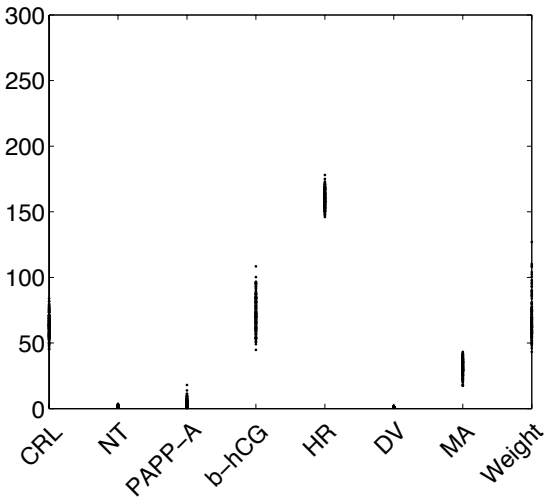
Example of ROC curves: Prenatal diagnosis of aneuploidy

Aneuploidy

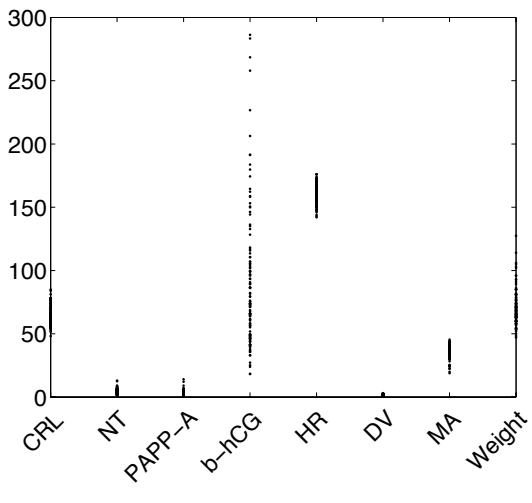
e.g. triploidy



Normal



Abnormal



Machine learning approach

→ Non-invasive diagnosis

→ Can signal risk for preeclampsia

→ Prevention of pregnancy complications