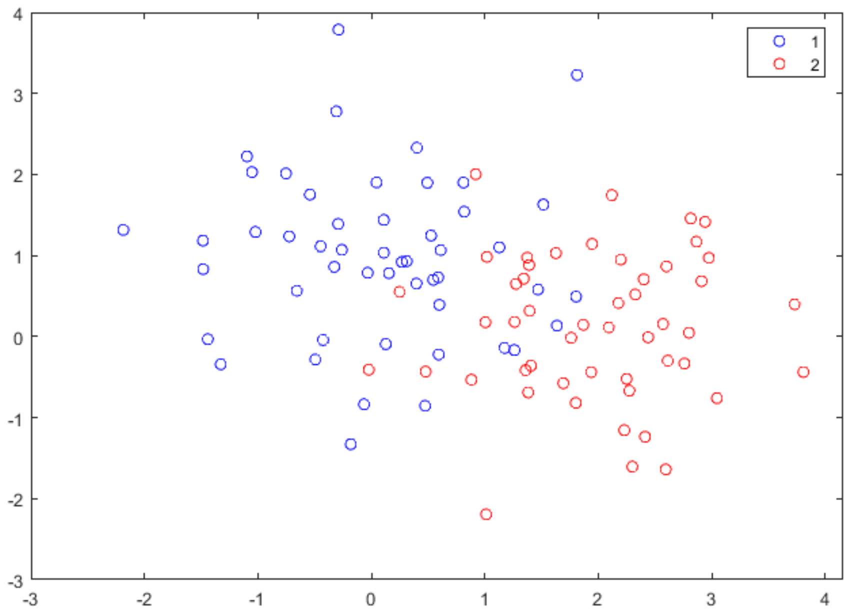
Classifiers

Laboratory 1. Bayes Classifiers

Example



Let’s assume we have a training dataset consisting observations belonging to two classes, 1 and 2 (50 data points for each class), with two features.

According to the formula, the decision rule takes the form of inequality

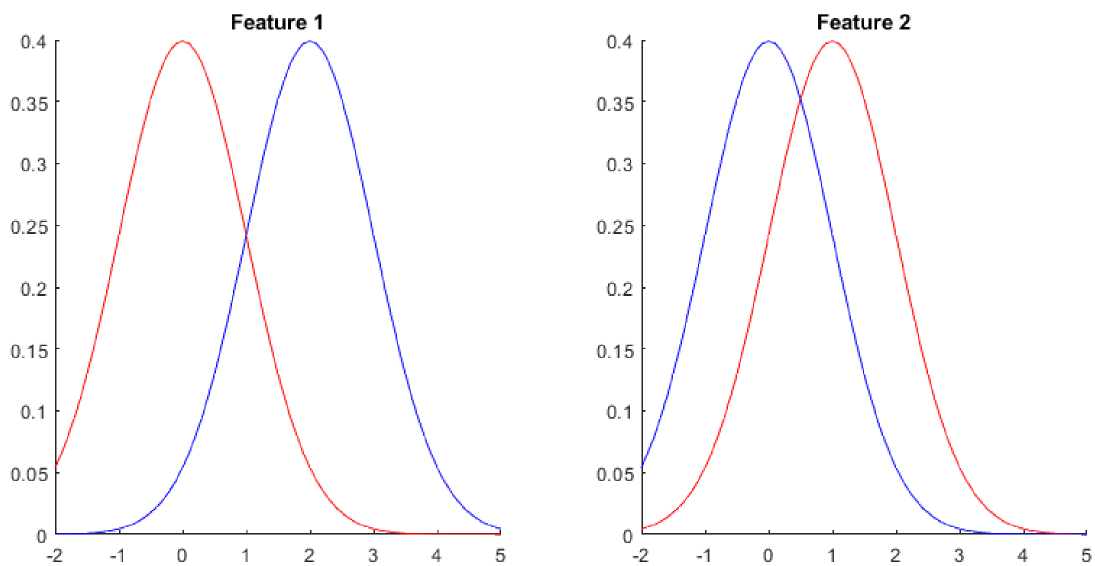
$$\frac{P(x\,|\,\omega_1)}{P(x\,|\,\omega_2)} > \frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \cdot \frac{P(\omega_2)}{P(\omega_1)}$$

$P(\omega_1), P(\omega_2)$ are the *a priori* probabilities, connected with numbers of observations of each class in the training dataset. 50 out of 100 total observations belong to class 1, so $P(\omega_1)=50/100=0.5$, same with the other class. λ_{ij} are the parameters denoting loss incurred for taking each decision. The simplest loss matrix is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, where the correct decision bears no loss, and the loss values for both incorrect decisions are equal. Using this loss matrix, the right side of our decision rule is constant and equal to

$P(\omega_2)/P(\omega_1)$, which equals 0.5/0.5 = 1.

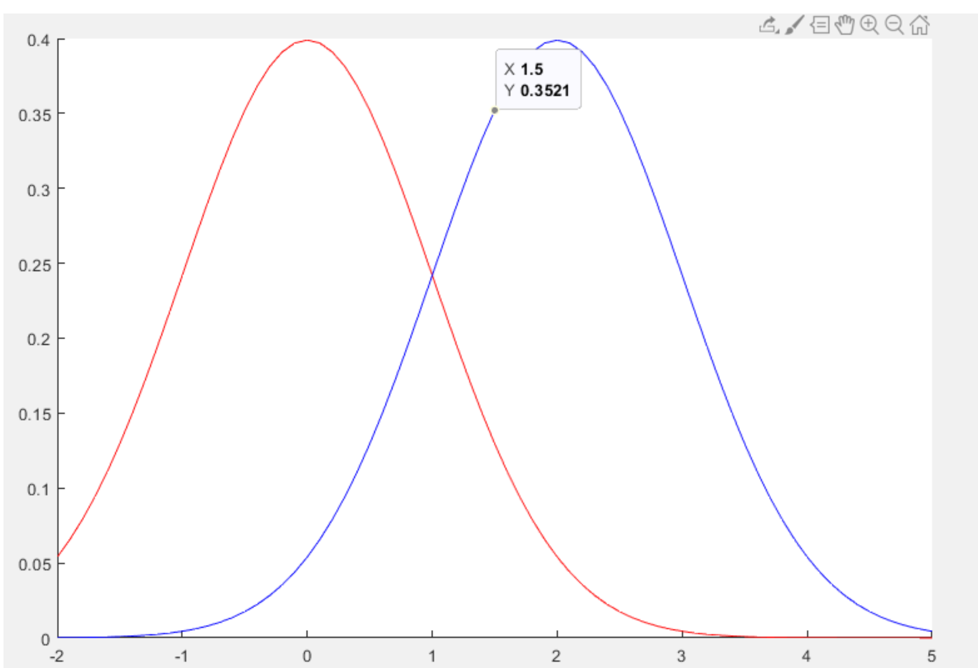
Now, suppose we want to classify a new observation with values [1.5,2.5].

We use the training set to estimate the probability density functions for each class and every feature:



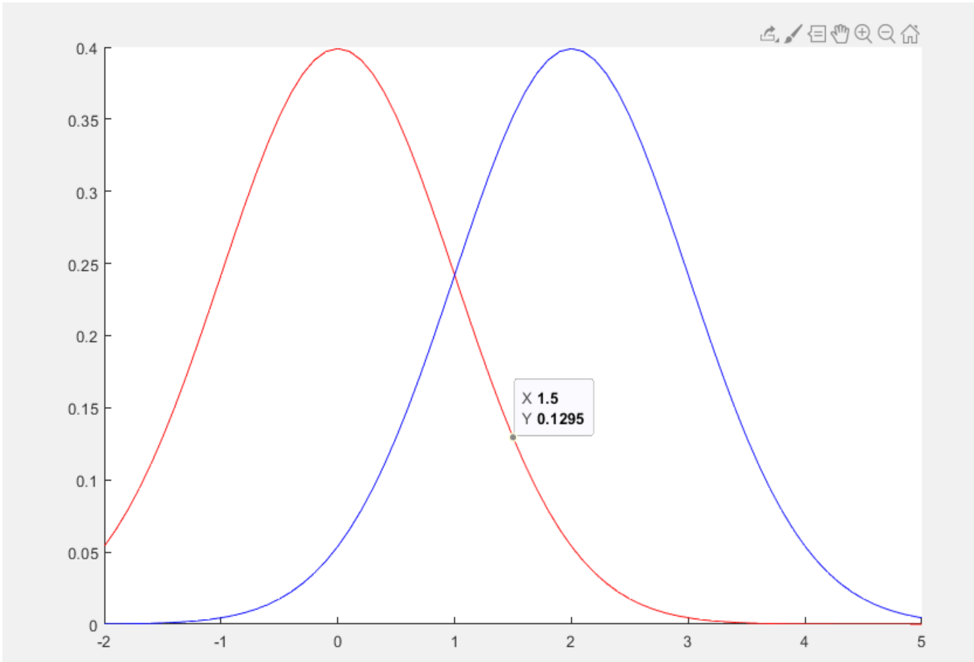
We can compute the conditional probabilities for the new data point. For the first feature:

$P(1.5|\omega_1)$ is the probability of obtaining value 1.5 given that the observation belongs to class 1



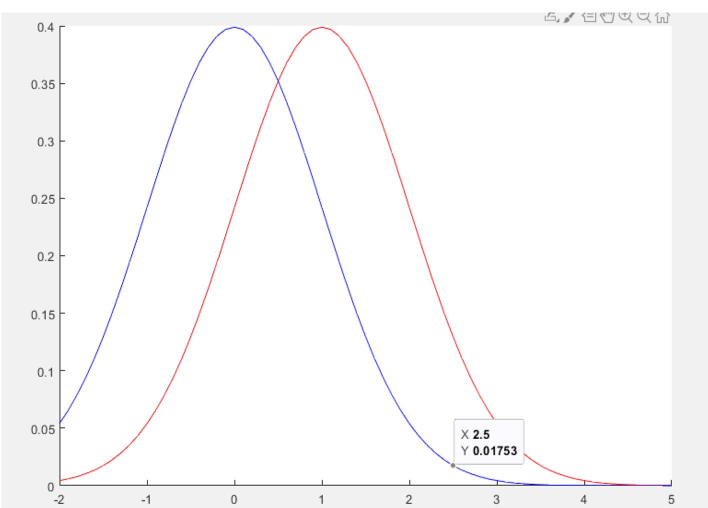
$P(1.5|\omega_1)=0.3521$

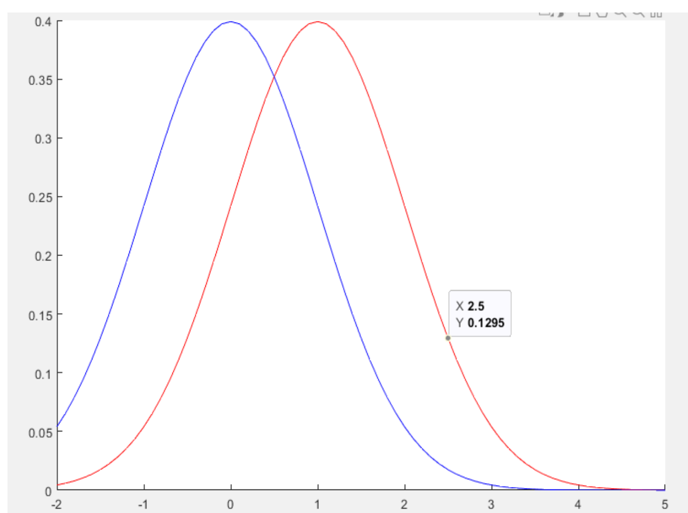
Similarily,



$P(1.5|\omega_2)=0.1295$

For the second feature:





$P(2.5|\omega_1)=0.0175$ $\qquad$ $P(2.5|\omega_2)=0.1295$

Because the classifier is “naive” we assume that the features are independent, so the probabilities for each feature are multiplied.

Therefore, the probability of values [1.5,2.5] under the condition that the observation belongs to class 1 is equal to:

$$P([1.5,2.5]|\omega_1)=0.3521*0.0175=0.0062$$

and for the second class:

$$P([1.5,2.5]|\omega_2)=0.1295*0.1295=0.0168$$

Going back to the classification rule,

$$\frac{P([1.5,2.5]|\omega_1)}{P([1.5,2.5]|\omega_2)}<1$$

which means the observation will be assigned to class 2.