1) Question 1

Solutions:

a) The number of features, the number of classes and their probabilities P(Ci)

We have 2 features x and y. We have 2 classes 0 and 1.

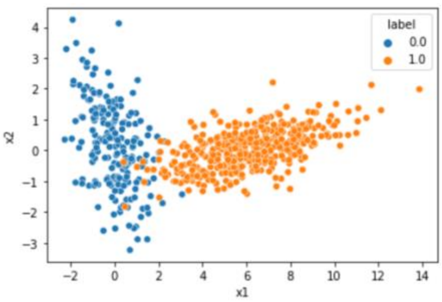
For find the probabilities we use the following formula:

$$P(C_i) = \frac{\sum C_i}{N}$$

$P(C_0) = \frac{\sum C_0}{N} = \frac{200}{600} = 0.333$

$P(C_1) = \frac{\sum C_1}{N} = \frac{400}{600} = 0.666$

b) Plot the data in a scatter plot using different colors for each class.



c) Partition the data into training (80%) and evaluation set (20%) randomly.

Our new shapes are the following:

Training □ ((480,2))

Evaluation □((120,2))

d) Compute the mean vector and covariance matrix for each class from the training set.

Mean Vector Formula:

$\boldsymbol{m_i} = \frac{1}{N_i}\sum_{j=1}^{N_i} \boldsymbol{x_{i,j}}$ where $x_{i,j}$ is sample j from class i, and Ni is the number of training samples from class i

Using this formula we find the mean vector the following:

	$\boldsymbol{m_0}$	$\boldsymbol{m_1}$
x	-0.020470	6.059459
y	0.14574475	0.02600738

Covariance Matrix S Formula:

$$\boldsymbol{S}: \boldsymbol{s_{ij}} = \frac{\sum_{t=1}^{N}(\boldsymbol{x_i^t} - \boldsymbol{m_i})(\boldsymbol{x_j^t} - \boldsymbol{m_j})}{\boldsymbol{N}}$$

Using this formula we find the covariance matrix for each class the following:

Class 0 covariance matrix:

$$s_0 = \begin{bmatrix} 0.85136583 & -0.57286741 \\ -0.57286741 & 1.8883089 \end{bmatrix}$$

Class 1 covariance matrix:

$$s_1 = \begin{bmatrix} 4.55393966 & 0.80551722 \\ 0.80551722 & 0.41910882 \end{bmatrix}$$

2) Question 2

If we design a linear discriminant classifier with a shared covariance matrix using MLE estimators;

Firstly, we create a shared covariance matrix the following formula:

$$S = \sum_i \hat{P}(C_i) S_i$$

For using this formula we find the estimated probabilities on training data for each class.

$$\hat{P}(C_0) = 0.3375$$
$$\hat{P}(C_1) = 0.6625$$

We found our Shared Covariance Matrix S as follows:

$$\text{S} = \begin{bmatrix} 3.30432099 & 0.34031241 \\ 0.34031241 & 0.91496385 \end{bmatrix}$$

We create a discriminant function the following formula:

$$g_i(x) = -0.5 * (x - m_i)^T S^{-1}(x - m_i) + log\, \hat{P}(C_i)$$

Which is a linear discriminant:

$$g_i(x) = {w_i}^T x + w_{i0}$$

Where:

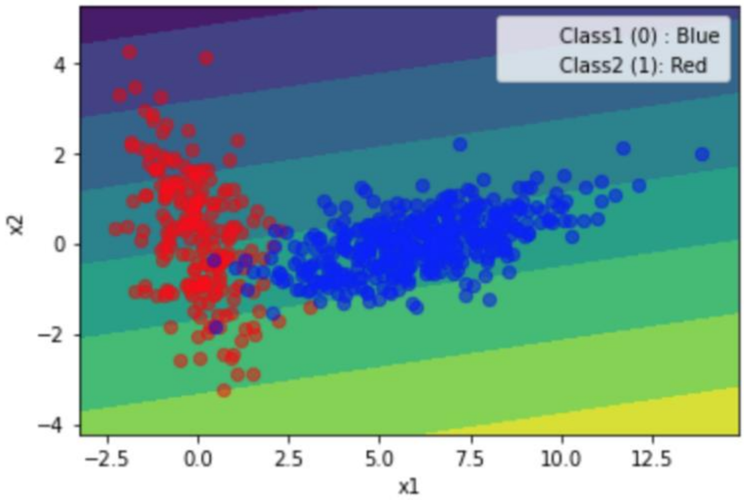
$$w_i = S^{-1} m_i\,, \quad w_{i0} = -0.5 * {m_i}^T S^{-1} m_i + log\, \hat{P}(C_i)$$

We calculate the cost using this formula and find accuracies for training and evaluation set:

Accuracy score for training data: 0.975

Accuracy score for evaluation data: 0.9583

Contour Plot for Linear Discriminant Function



3) Question 3

If we want to design a quadratic discriminant classifier using different covariance matrices:

Firstly we have to calculate covariance matrices for each class.

We calculated our covariance matrices in the first question, item 4.

Class 0 covariance matrix:

$$s_0 = \begin{bmatrix} 0.85136583 & -0.57286741 \\ -0.57286741 & 1.8883089 \end{bmatrix}$$

Class 1 covariance matrix:

$$s_1 = \begin{bmatrix} 4.55393966 & 0.80551722 \\ 0.80551722 & 0.41910882 \end{bmatrix}$$

We create a quadratic discriminant function using the following formula:

$$g_i(x) = -0.5 * log|S_i| - 0.5 * (x^T {S_i}^{-1} x - 2x^T {S_i}^{-1} m_i + {m_i}^T {S_i}^{-1} m_i) + log\, \hat{P}(C_i)$$

$$g_i(x) = x^T W_i x + {w_i}^T x + w_{i0}$$

Where:

$$W_i = -0.5 * {S_i}^{-1}, w_i = S^{-1} m_i\, , \quad w_{i0} = -0.5 * {m_i}^T S^{-1} m_i - 0.5 * log|S_i| + log\, \hat{P}(C_i)$$

We calculate the cost using this formula and find accuracies for training and evaluation set:

Accuracy score for training data: 0.9875

Accuracy score for evaluation data: 0.9583

Contour Plot for Quadratic Discriminant Function