Istanbul Technical University- Spring 2017
BLG527E Machine Learning
Homework 1

Omid Abdollahi Aghdam
student number: 504151520
abdollahi15@itu.edu.tr
abdollahi.omid@gmail.com

March 6, 2017

Q1) The code for Q1a and Q1b is written in file *'q1.py'* and running the code will plot the histogram of means for all four cases, as you can see in figure 1, and will save an image file with the name *'hists.png'*.
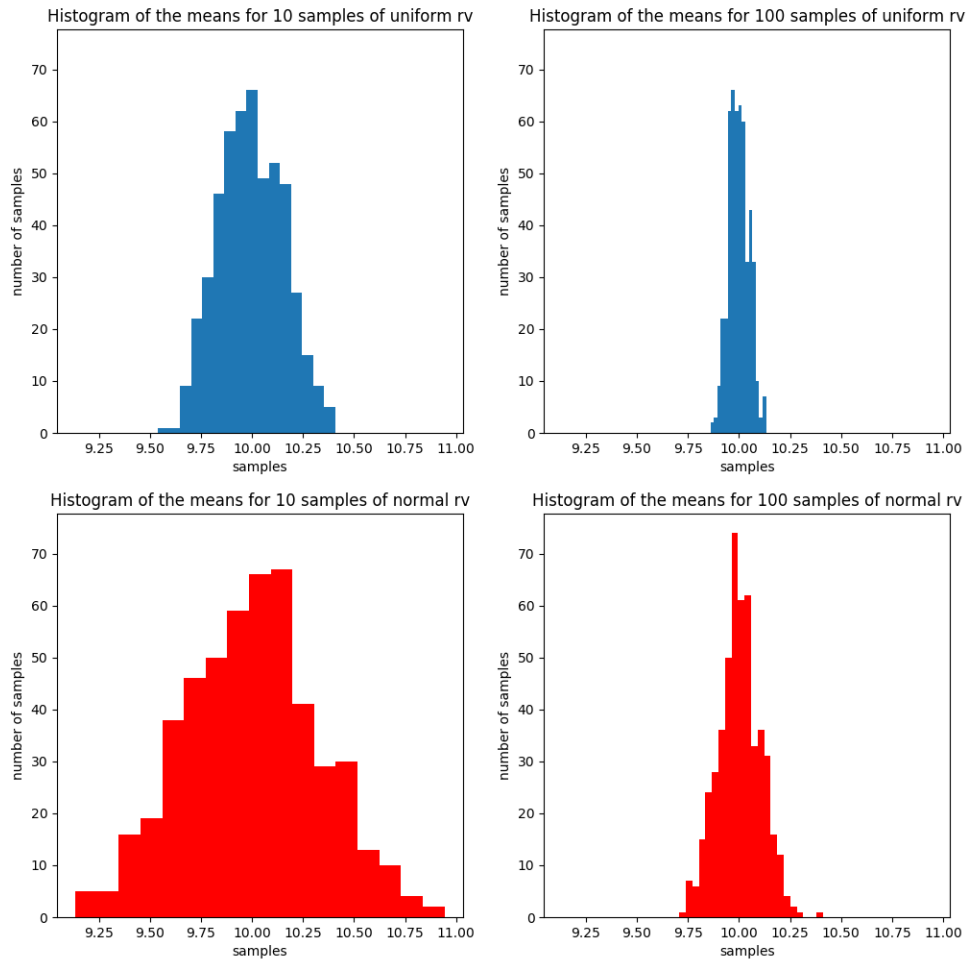

Figure 1: Histograms of means.

Q1c) According to Central Limit Theorem, if we take N sample from any distribution with mean μ and variance σ^2, calculate the mean of samples, and repeat this experiments several times, in this homework we repeat 500 times. As a result, the distribution of the mean get closer to a Normal Distribution with mean μ and variance σ^2/N as the number of samples N increases.

Similarities:

1. All four plots looks like normal distribution, and as the number of samples increases to 100 the histograms become more similar to Normal distribution pdf.

2. The mean of all histograms are $\mu = 10$ because the mean of distributions which we draw random samples from equal 10.

3. The variance of sampling distribution of means get closer to each other as N increases, so, the variance of histograms for $N = 100$ are closer to each other than histograms for $N = 10$.

Differences:

1. The only difference I can see is the variances. However as the N increases the differences become less significant.

Q2) a) We have the discriminant function as follow:

$$g_i = \ln(p(x|w_i)) + \ln(p(w_i)) \tag{1}$$

Also we have:

$$x \in \Re$$
$$class1 : x \sim \mathbf{N}(\mu_1, \sigma)$$
$$class2 : x \sim \mathbf{N}(\mu_2, \sigma)$$
$$p(w_1) = p(w_2)$$

Using the information above and the following equation we can derive discriminant functions for class 1 and class 2.

$$p(x|w_i) = \frac{1}{\sqrt{2\pi}\sigma_i} exp\left[-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right] \tag{2}$$

Now we plug in the equation 2 in equation 1 and derive discriminant functions.

$$g_1(x) = \ln\left[\frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(x - \mu_1)^2}{2\sigma^2}\right]\right] + \ln(p(w_1)) \tag{3}$$

We use logarithm properties to derive the discriminant function for class 1, $g_1(x)$, and later we just replace μ_1 with μ_2 to get the discriminant function for class 2, $g_2(x)$.

$$g_1(x) = \ln(1) - \frac{1}{2}ln(2\pi) - ln(\sigma) - \frac{(x - \mu_1)^2}{2\sigma^2}ln(e) + ln(p(w_1))$$
$$g_1(x) = -\frac{1}{2}ln(2\pi) - ln(\sigma) - \frac{(x - \mu_1)^2}{2\sigma^2} + ln(p(w_1))$$

The discriminant functions for class 1 and class 2 are:

$$g_1(x) = -\frac{1}{2}ln(2\pi) - ln(\sigma_1) - \frac{(x - \mu_1)^2}{2\sigma_1^2} + ln(p(w_1)) \tag{4}$$

3

$$g_2(x) = -\frac{1}{2}ln(2\pi) - ln(\sigma_2) - \frac{(x-\mu_2)^2}{2\sigma 2^2} + ln(p(w_2)) \qquad (5)$$

b) To find the separation surface we set discriminant equations equal.

$$g_1(x) = g_2(x) \qquad (6)$$

Now we replace equation 4 and equation 5 in equation 6 and find the equation for separating surface.

$$-\frac{1}{2}ln(2\pi) - ln(\sigma_1) - \frac{(x-\mu_1)^2}{2\sigma_1^2} + ln(p(w_1)) = -\frac{1}{2}ln(2\pi) - ln(\sigma_2) - \frac{(x-\mu_2)^2}{2\sigma_2^2} + ln(p(w_2))$$

$$\left(\frac{1}{2\sigma_2^2} - \frac{1}{2\sigma_1^2}\right)x^2 + \left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right)x + \frac{\mu_2^2}{2\sigma_2^2} - \frac{\mu_1^2}{2\sigma_1^2} + \ln\frac{\sigma_2}{\sigma_1} + \ln\frac{p(w_1)}{p(w_2)} = 0 \qquad (7)$$

As you can see separating surface function is a quadratic function $(ax^2 + bx + c = 0)$ and we can find its roots using the following formula.

$$x1, x2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In this example $\sigma_1 = \sigma_2$, so $a = 0$ so we can eliminate quadratic term. The new separating surface equation can be written as follow:

$$x = \frac{2\sigma^2 \ln\left(\frac{p(w_2)}{p(w_1)}\right) + \mu_1^2 - \mu_2^2}{2(\mu_1 - \mu_2)} \qquad (8)$$

4

c) We plug $\mu_1 = 5$, $\mu_2 = 15$, $\sigma = 5$ and $p(w_1) = p(w_2) = 0.5$ in equation 8. As a result we have the following vertical line as the separating surface. You can see the plot in figure 2. Separating line specifies which class the inputs belong to. The code is written in *'pdfs.py'*.
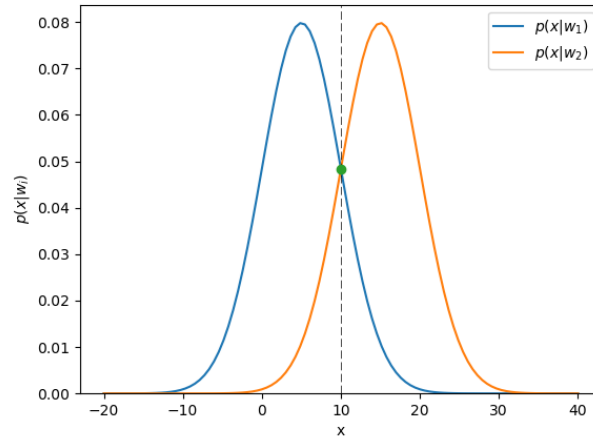
$$x = 10$$

Figure 2: pdf's of the two classes and separating surface, $x = 10$.

d) We plug the new class probabilities $p(w_1) = 0.8$ and $p(w_2) = 0.2$ in the equation 8 and the new separating surface is **x=13.4657**.The new plot is shown in figure 3.
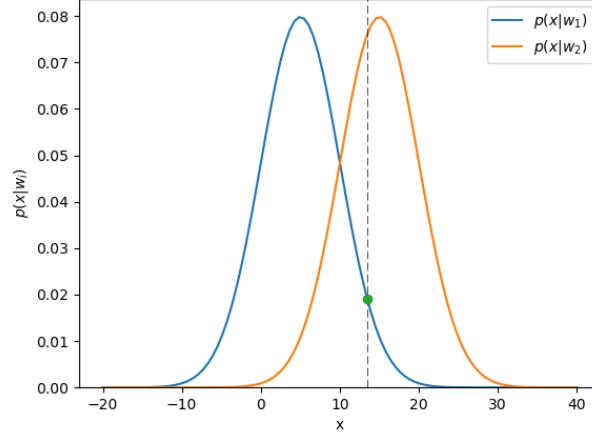

Figure 3: pdf's of the two classes and separating surface, $x = 13.4657$.

e) As $p(w_2) = p(w_1)$ in **c)** so we generate equal number of samples for datasets. We generated 200 samples from each distribution. Finally, we use the mean and standard deviation of generated datasets and plug them in equation 7 to calculate the separating point. The equation has two results we select the result that is in range of the generated x values. The histograms are shown in figure 4.
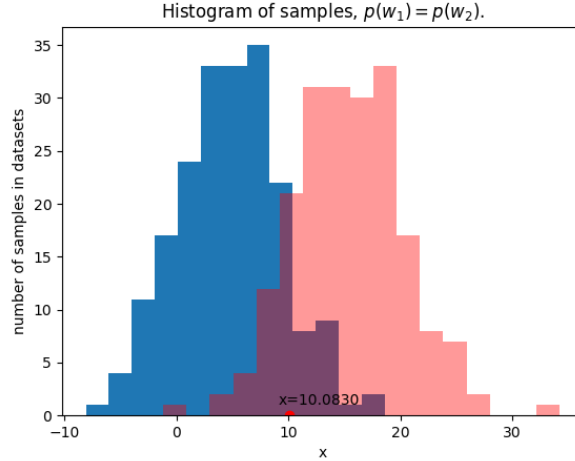

Figure 4: The histograms of generated datasets .

For **d)** we generated 320 samples from class 1 and 80 samples from class 2. We calculate the separating point using equation 7 and select the x that is in the range of generated x values. The histograms for the generated datasets are shown in figure 5.
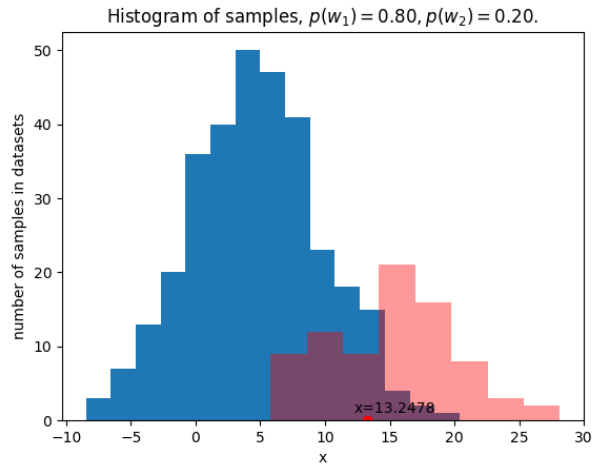

Figure 5: The histograms of generated datasets .

Source codes for the first question is written in file *'q1.py'* and the codes for the second question is written in file *'q2.py'*. In order to run codes successfully *numpy, matplotlib, scipy* python libraries must have been installed in your system.