EE559

Homework 7

Sarthak Maharana

Q1.

(a) The given class conditional probabilities are:

$P(x/S_i) = N(\underline{x}, \underline{m_i}, \underline{\underline{\Sigma_i}})$, $i = 1,2$.

It's a 2-class problem.

$P(S_1) = P(S_2) = 0.5$ and $\underline{m_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $m_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

$\underline{\underline{\Sigma_1}} = \underline{\underline{\Sigma_2}} = \begin{bmatrix} 2 & 3 \\ 3 & 6.5 \end{bmatrix}$

We know, $P(\text{error}) = 1 - P(\text{correct})$

$$= 1 - \sum_{i=1}^{C} P(S_i) \int_{\Gamma_i} P(\underline{x}/S_i) d\underline{x}$$

Hence, to minimise the error, the integral over all the classes has to be maximised.

According to Bayes' decision,

$P(x/S_i)\,P(S_i) > P(x/S_j)\,P(S_j) \quad \forall\, i \neq j$, then $\underline{x} \in \Gamma_i$

$\therefore\ P(x/S_1)\,P(S_1) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} P(x/S_2)\,P(S_2)$.

or $$\frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{\frac{-1}{2}(\underline{x}-\underline{m_i})^T \underline{\underline{\Sigma_i}}^{-1}(x-\underline{m_i})\right\} P(S_1) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}}$$

$$\frac{1}{(2\pi)^{D/2}(\Sigma_j)^{1/2}} \exp\left\{-\frac{1}{2}(\underline{x}-\underline{m_j})^T \underline{\underline{\Sigma_2}}^{-1}(\underline{x}-m_2)\right\} P(S_2)$$

Since, $p(S_1) = p(S_2) = 0.5$, take logarithms on both the sides :

$$\tilde{g}_i(\underline{x}) = -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln(|\Sigma_i|) - \frac{1}{2}(x-m_i)^T \underline{\underline{\Sigma}}_i^{-1}(x-m_i)$$

Ignoring constant terms :

D = 2

$$-\frac{1}{2}(\underline{x}-\underline{m}_1)^T \underline{\underline{\Sigma}}_1^{-1}(\underline{x}-\underline{m}_1) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} -\frac{1}{2}(x-m_2)^T \underline{\underline{\Sigma}}_2^{-1}(x-m_2).$$

$$\text{Now, } \underline{\underline{\Sigma}}_1^{-1} = \underline{\underline{\Sigma}}_2^{-1} = \begin{bmatrix} 2 & 3 \\ 3 & 6.5 \end{bmatrix}^{-1} = \begin{bmatrix} 1.625 & -0.75 \\ -0.75 & 0.5 \end{bmatrix}.$$

$$\text{or } (\underline{x}-\underline{m}_2)^T \underline{\underline{\Sigma}}_2^{-1}(x-m_2) - (x-m_1)^T \underline{\underline{\Sigma}}_1^{-1}(x-m_1) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 0.$$

$$[x_1-2 \quad x_2+3]\begin{bmatrix} 1.625 & -0.75 \\ -0.75 & 0.5 \end{bmatrix}\begin{bmatrix} x_1-2 \\ x_2-3 \end{bmatrix} - [x_1-1 \quad x_2+1]\begin{bmatrix} 1.625 & -0.75 \\ -0.75 & 0.5 \end{bmatrix}\begin{bmatrix} x_1-1 \\ x_2+1 \end{bmatrix} \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 0$$

$$\Big[(x_1-2)1.625 - (x_2-3)0.75 \quad -(x_1-2)0.75 + (x_2-3)0.5\Big]_{1\times2}\begin{bmatrix} x_1-2 \\ x_2-3 \end{bmatrix}_{2\times1}$$

$$-\Big[(x_1-1)(1.625) - (x_2+1)0.75 \quad -(x_1-1)0.75 + (x_2+1)0.5\Big]\begin{bmatrix} x_1-1 \\ x_2+1 \end{bmatrix} \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 0$$

$$(x_1-2)^2 \cancel{1.625} - (x_2-3)(x_1-2)0.75 - (x_1-2)(x_2-3)0.75 + (x_2-3)^2 0.5$$

$$-(x-1)^2\cancel{1.625} + (x_1-1)(x_2+1)0.75 + (x_1-1)(x_2+1)0.75 - (x_2+1)^2 0.5 \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 0$$

On simplifying, (after rearranging) $2.75x_1 - 2.5x_2 \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 1.625$ // ← Decision rule.

(b) Now, $P(S_1) = 0.1$ and $P(S_2) = 0.9$.

$$P(\underline{x}|S_1)\,P(S_1) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} P(\underline{x}|S_2)\,P(S_2)$$

$$P(\underline{x}/S_1) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} P(\underline{x}|S_2)\,9.$$

Taking logarithms,

$$\ln(P(\underline{x}|S_1)) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} \ln(P(\underline{x}|S_2)) + \ln 9.$$

After ignoring the constant terms,

$$-\frac{1}{2}(\underline{x}-\underline{m_1})^T \underline{\underline{\Sigma}}_1^{-1}(\underline{x}-\underline{m_1}) \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} -\frac{1}{2}(\underline{x}-\underline{m_2})^T \underline{\underline{\Sigma}}_2^{-1}(\underline{x}-\underline{m_2}) + \ln 9.$$

$$\frac{1}{2}\left[(\underline{x}-\underline{m_2})\underline{\underline{\Sigma}}_2^{-1}(\underline{x}-\underline{m_2}) - (\underline{x}-\underline{m_1})^T\underline{\underline{\Sigma}}_1^{-1}(\underline{x}-\underline{m_1})\right] \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} \ln 9.$$

$$\frac{1}{2}\left[\begin{bmatrix} x_1-2 & x_2-3\end{bmatrix}\begin{bmatrix}1.625 & -0.175\\ -0.175 & 0.5\end{bmatrix}\begin{bmatrix}x_1-2\\ x_2-3\end{bmatrix} - \begin{bmatrix} x_1+1 & x_2+1\end{bmatrix}\begin{bmatrix}1.625 & -0.175\\ -0.175 & 0.5\end{bmatrix}\begin{bmatrix}x_1-2\\ x_2-3\end{bmatrix}\right] \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} \ln 9 = 2.1972$$

On simplifying,

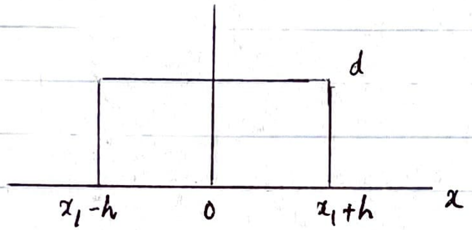
$$\text{or} \quad 2.75x_1 - 2.5x_2 - 1.625 \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 4.394$$

$$\text{or} \quad 2.75x_1 - 2.5x_2 \underset{\Gamma_2}{\overset{\Gamma_1}{\gtrless}} 6.019 \quad // \quad \leftarrow \text{Decision rule.}$$

PLOTS HAVE BEEN ATTACHED LATER !!

2. Target : KDE with a binary window

(a) $\phi\left(\frac{u}{h}\right) = \left[\left[-1 \le \frac{u}{h} \le 1\right]\right]$



(Rectangular window)

$$\therefore p(x_1 | s) = \begin{cases} d & x \in (x_1 - h,\ x_1 + h) \\ 0 & \text{else.} \end{cases}$$

$$2hd = \frac{1}{\text{length (class vector)}} \quad \text{or} \quad d = \frac{1}{2h(\#\text{ of points in class } S_i)}$$

(b) Given,

S_1 : 0, 0.4, 0.9, 1.0, 6.0, 8.0
S_2 : 2.0, 4.0, 4.5, 5.0, 5.8, 6.7, 7.0.

$$P(S_1) = \frac{\#\text{ of elements in } S_1}{\#\text{ of elements in } (S_1+S_2)} = \frac{6}{13}.$$

$$P(S_2) = \frac{\#\text{ of elements in } S_2}{\#\text{ of elements in } (S_1+S_2)} = \frac{7}{13}. //$$

(c) (d) (e) ALL THE PLOTS HAVE BEEN ATTACHED AT THE END !!

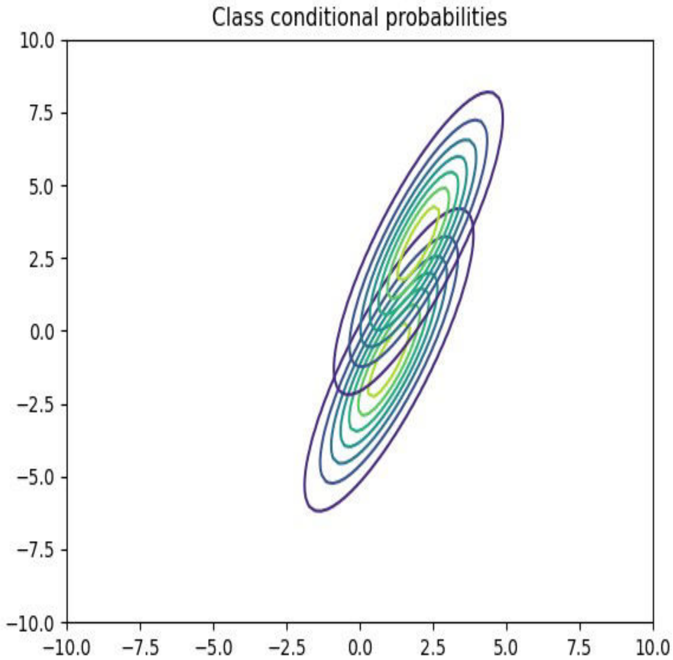
3. DONE LATER !!

(AT THE END)

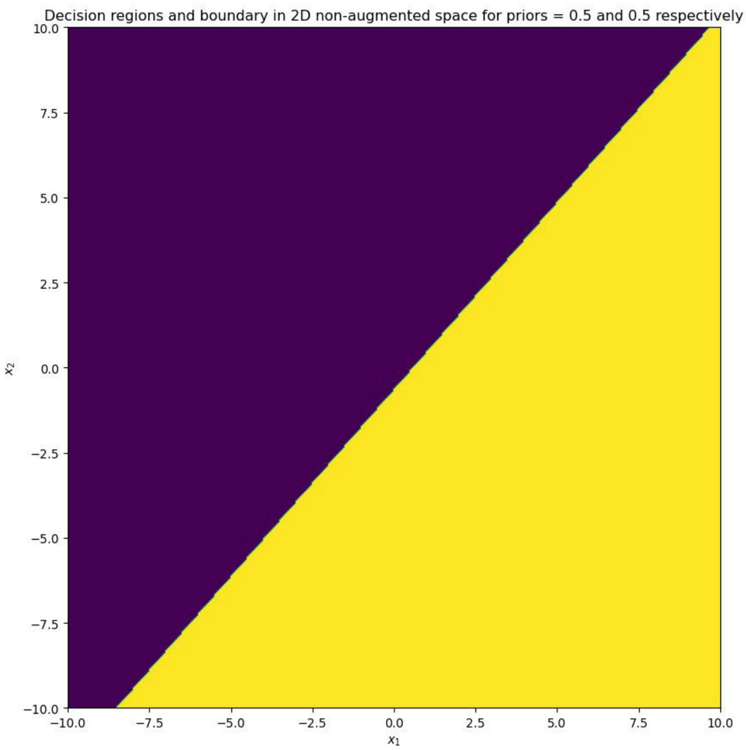
HOMEWORK 7

PROBLEM 1

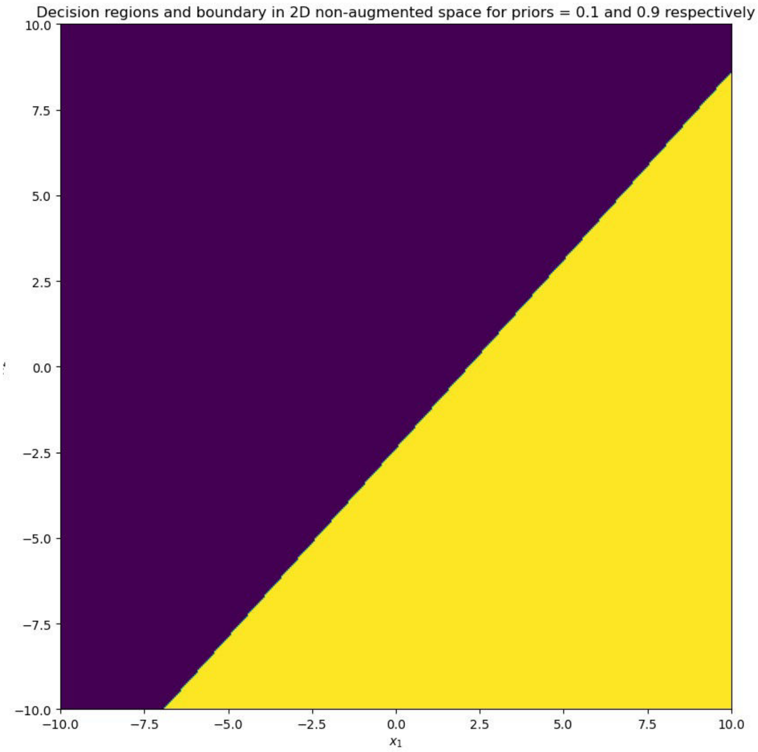
(a) Since $P(S_1) = P(S_2) = 0.5$ and with a non-diagonal covariance matrix, as expected we **obtained two ellipsoids with shifted axes**. The generated class probability distributions are as follows :



The figure below is an illustration of the decision boundary regions obtained by the Bayes' minimum error classifier when **$P(S_1) = P(S_2) = 0.5$**. **Here, the region marked yellow denotes class S_1 and the region marked purple denotes class S_2, since we use scipy.stats.multivariate_normal to generate the probability distribution.**

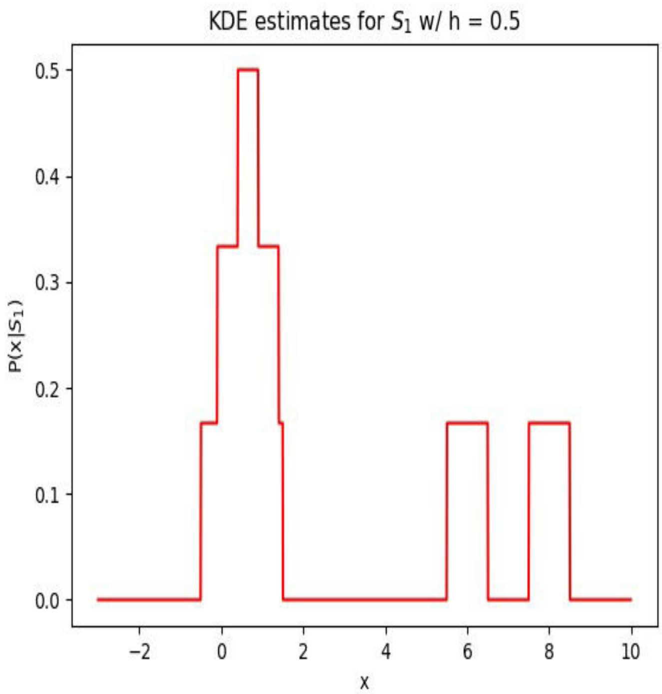


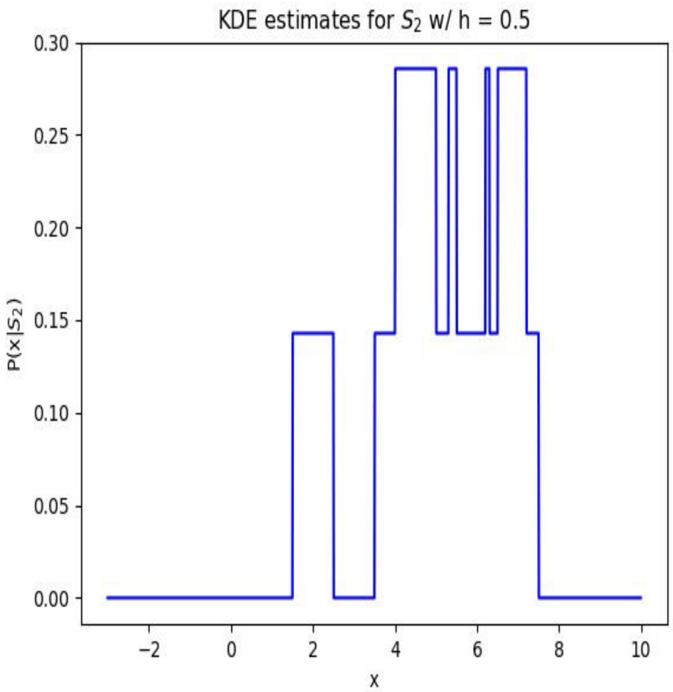
(c) The figure below is an illustration of the decision boundary regions obtained by the Bayes' minimum error classifier when **$P(S_1) = 0.1$ and $P(S_2) = 0.9$**. Here, the r**egion marked yellow denotes class S_1 and the region marked purple, denotes class S_2, since we use scipy.stats.multivariate_normal to generate the probability distribution. Clearly, we see a shift in the linear decision boundary.**



PROBLEM 2

(a) The plots of the KDE estimates are as follows (**h = 0.5**) :

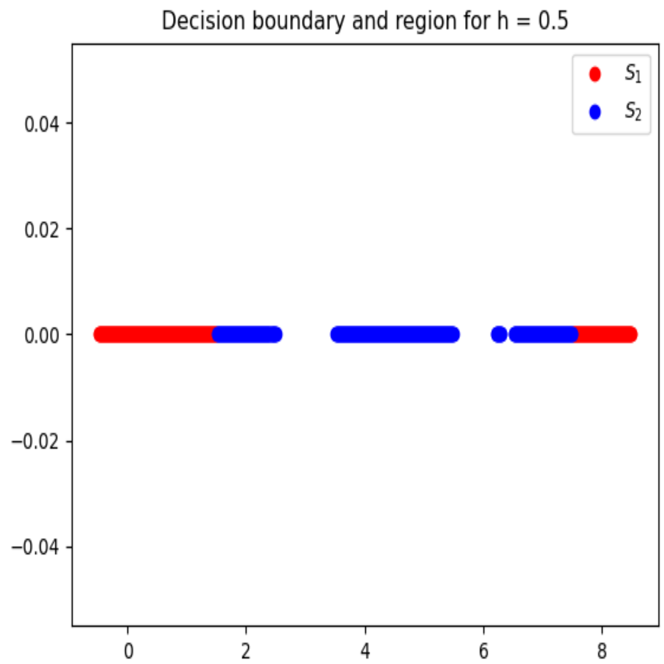




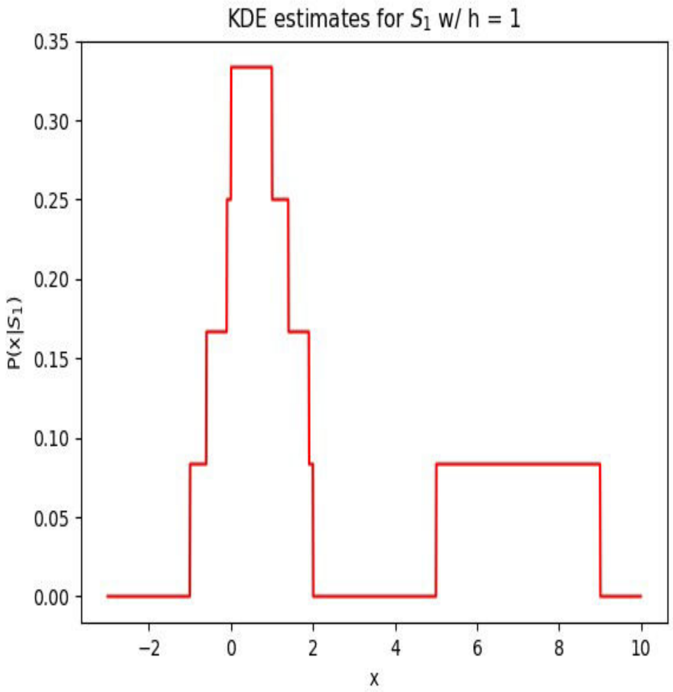
(b) Already done (handwritten).

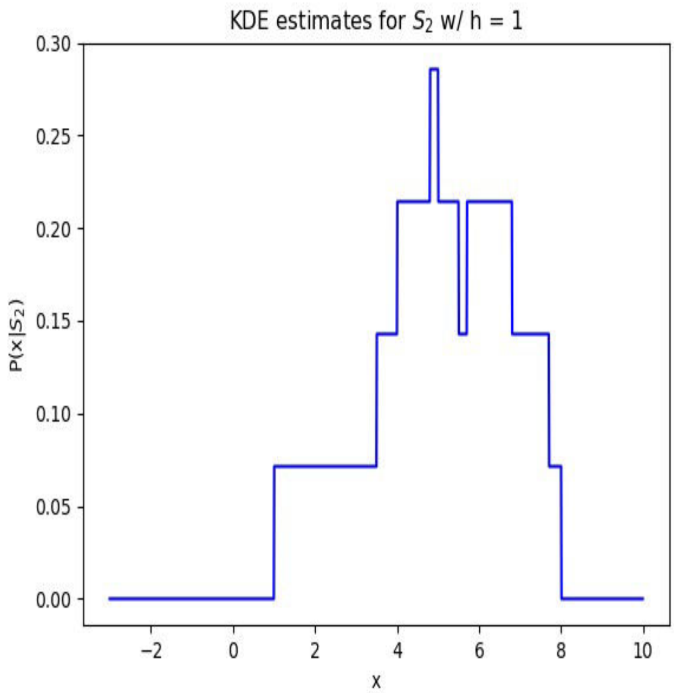
$P(S_1) = 6/13$
$P(S_2) = 7/13$

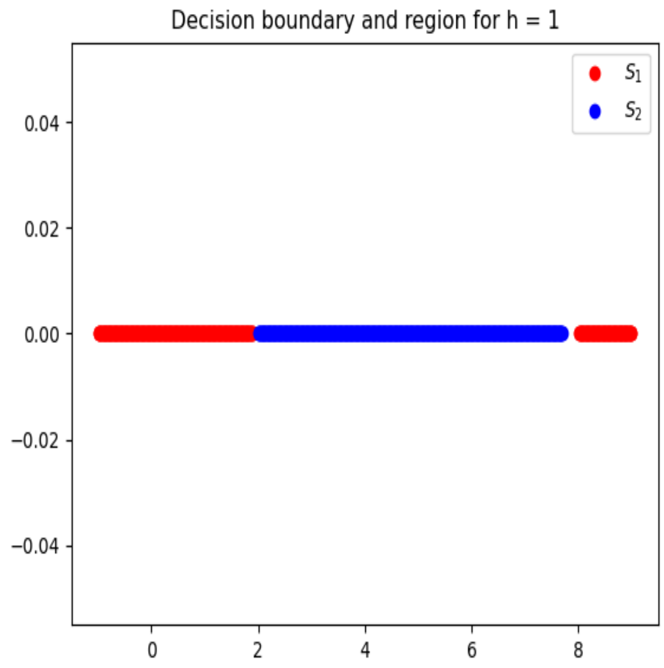
(c) The decision boundaries and regions for a Bayes minimum-error classifier, based on KDE:



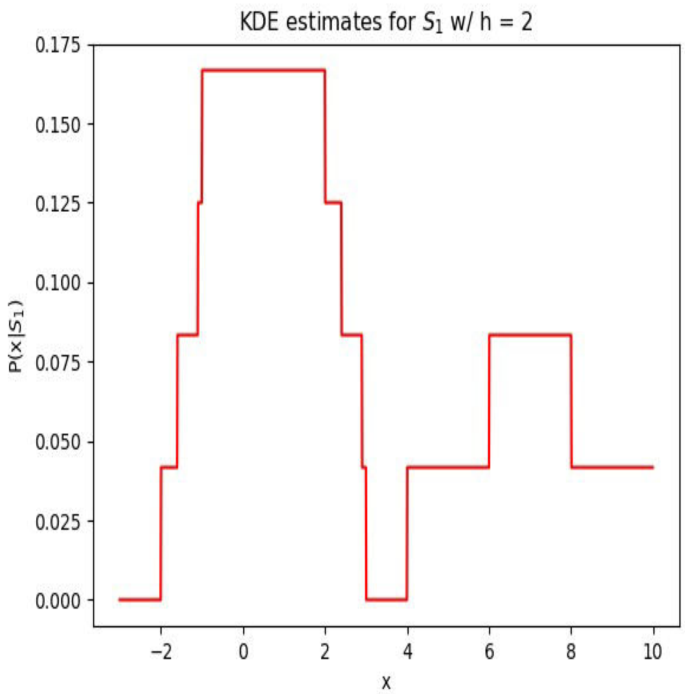
(d) The next three figures illustrate the KDE estimates and the decision boundary and regions for a width **h of 1** for classes S_1 and S_2 :

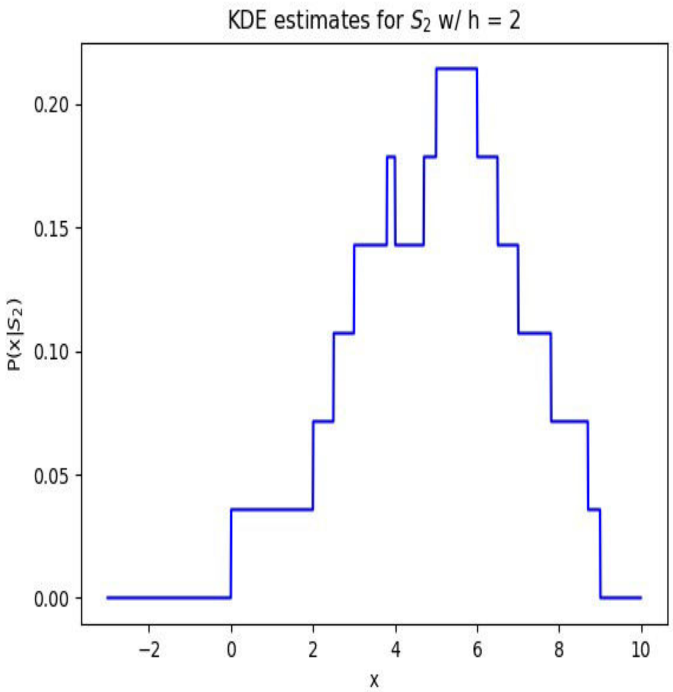


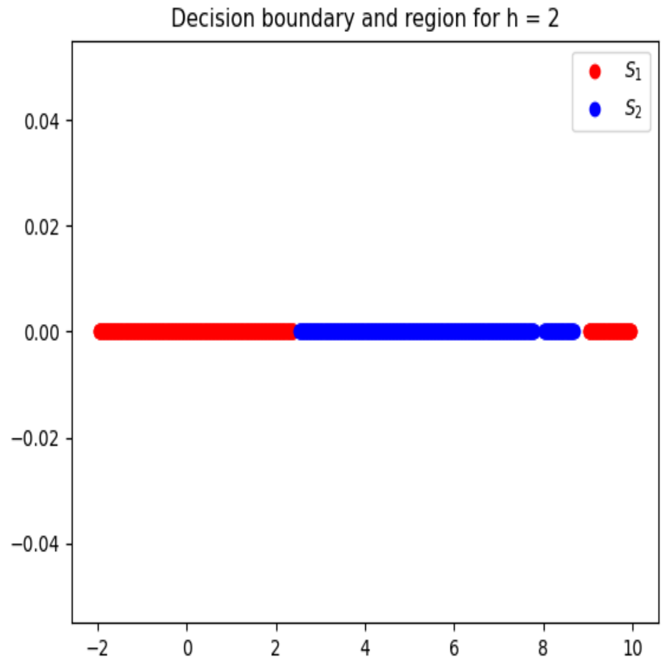




(e) The next three figures illustrate the KDE estimates and the decision boundary and regions for a width **h of 2** for classes S_1 and S_2 :







PROBLEM 3

(a) For the following test points (x = 0, 0.4), the y'(x) = **0.05 and 0.21** respectively.

(b) The MSE on the "test set" predictions of (a) = **0.0025.**

(c) Using the defined kernel function, the predictions for **x = 0, 0.4 are 0.0367 and 0.1967 respectively.** Since we use a weighted KNN to perform regression, we expect the MSE between the predictions and target to be lower. And yes, on doing this, the MSE = **0.0013** on the “test set”.