Universidad de Buenos Aires
Facultad de Ingenieria
Año 2018 - 2er Cuatrimestre

Aprendizaje Estadístico, Teoría y aplicación

TAREA

ALUMNOS:

Cabrera, Mauricio Luca - #101334
mlcabrera@fi.uba.ar
Flores Rodríguez, José Julián - Maestría en Ingeniería Matemática
julian12310@gmail.com
Cajachuán, Kevin - #98725
kevincajachuan@hotmail.com

Índice

1. Clase 1

1.1. Ejercicio 1

Sea una función convexa $f(x)$ no negativa, x_0,x_1,x_2 tales que :

$$x_1 \leq x_0 \leq x_2$$

$$x_0 = p_1 x_1 + p_2 x$$

$$p_1 \geq 0, p_2 \geq 0, p_1 + p_2 = 1$$

Y sea una recta $g(x)$ tal que:

$$g(x_0) = f(x_0)$$

$$g(x_1) = f(x_1)$$

Probar la siguiente desigualdad:

$$f(x_0) \leq p_1 f(x_1) + p_2 f(x_2)$$

La solucion de este problema es simplemente plantear $g(x) = ax + b$ y que como la función es $f(x)$ es convexa en el intervalo $[x_1, x_1]$ vale lo siguiente.

$g(x_1) = ax_1 + b = f(x_1)$ y $g(x_2) = ax_2 + b = f(x_2)$ entonces $g(x_0) = g(p_1 x_1 + p_2 x_2) = a(p_1 x_1 + p_2 x_2) + b = ap_1 x_1 + ap_2 x_2 + b$ y como $p_1 + P_2 = 1$ puedo multiplicar al termino b por $(p_1 + P_2)$ sin alterar el resultado.

Esto resulta en que $p_1 x_1 + ap_2 x_2 + b(p_1 + P_2)$ con lo que queda que

$$g(x_0) = p_1(ax_1 + b) + p_2(ax_2 + b) = p_1 f(x_1) + p_2 f(x_2)$$

Ahora como $f(x) \leq g(x) \forall x \in [x_1, x_2]$, vale la desigualdad:

$$f(p_1 x_1 + p_2) \leq p_1 f(x_1) + p_2 f(x_2)$$

1.2. Ejercicio 2 (opcional)

Utilizando $E[f(x)m(x)] = E[Yf(x)]$, y algún artilugio que crea conveniente, demostrar que vale la siguiente ecuacion:

$$E[\phi(x)Y|X] = \phi(x)E[Y|X]$$

1.3. Ejercicio 3: Simulacion de 200 puntos

Los pares ordenados fueron simulados con las siguientes condiciones:

- $X \sim \mathcal{N}(0,1)$ una variable aleatoria truncada en el intervalo $[-1,1]$.
- la funcion de regresión esta dado por la siguiente forma:

$$E[Y|_{X=x}] = m(x) = \begin{cases} \frac{(x+2)^2}{2} & \text{si } -1 \leq x < -0,5 \\ \frac{x}{2} + 0,875 & \text{si } -0,5 \leq x \leq 0 \\ -5(x-0,2)^2 + 1,075 & \text{si } 0 < x < 0,5 \\ x + 0,125 & \text{si } 0,5 \leq x < 1 \end{cases}$$

- $(Y - m(X)) \sim \mathcal{N}(0, \delta^2(X))$ con $\delta(X) = 0{,}2 - 0{,}1\cos(2\pi X)$

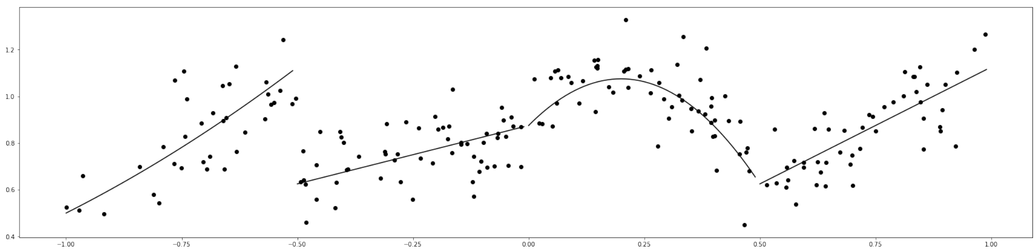


Figura 1: 200 puntos simulados junto con la funcion de regresion.

1.4. Ejercicio 4: no entregable

Dado el problema de decisión introducido en el "Diseño del receptor de una comunicación binaria", verificar que:

$$\delta(r) = \mathbb{1}\{P(S=1|R=r) > P(S=0|R=r)\}$$

2. Clase 2

2.1. Ejercicio 5

Resolver la siguiente cuenta:

$$P(X,Y) = \int_0^{1/2} (1-x)\ dx + \int_{1/2}^{3/4} x\ dx$$

$$= \left(x - \frac{x^2}{2} \right) \Big|_0^{1/2} + \frac{x^2}{2} \Big|_{1/2}^{3/4}$$

$$= \left(\frac{1}{2} - \frac{(1/2)^2}{2} \right) + \left(\frac{(3/4)^2}{2} - \frac{(1/2)^2}{2} \right)$$

Por lo tanto:

$$P(X, Y) = \frac{17}{32}$$

3. Clase 3

3.1. Ejercicio 8: Simulacion de Seleccion

Abajo se simulara distantas cantidades de puntos de color y los cuales se clasificaran con distintos colores de la siguiente forma:

- Se dibujan N puntos de colores, con la probabilidad de $1/2$ de que sea azul y $1/2$ de que sea rojo.
- Los rojos tendran una distribucion uniforme cubriendo un triangulo isosceles de base la recta $(0, 0) \to (1, 0)$ y de altura 1
- Los azules tendran una distribucion uniforme cubriendo un triangulo isosceles de base la recta $(0, 1) \to (1, 1)$ y de altura -1
- Al terminar de pintar los puntos, usando la regla del histograma, se crearan una cuadrilla que posteriorente se pintaran de un color dependiendo de cuantos colores entraron ahi.

Para esta simulacion uso $h_n = \frac{1}{\sqrt[3]{n}}$ y se uso el siguiente criterio para la clasicacion:

$$g_n(X) = \begin{cases} 1 & \text{si } \sum_{i=0}^{n} \mathbb{1}\{Y = 1\} \mathbb{1}\{X_i \in A(x)\} \geq \sum_{i=0}^{n} \mathbb{1}\{Y = 0\} \mathbb{1}\{X_i \in A(x)\} \\ 0 & \text{e.o.c} \end{cases}$$

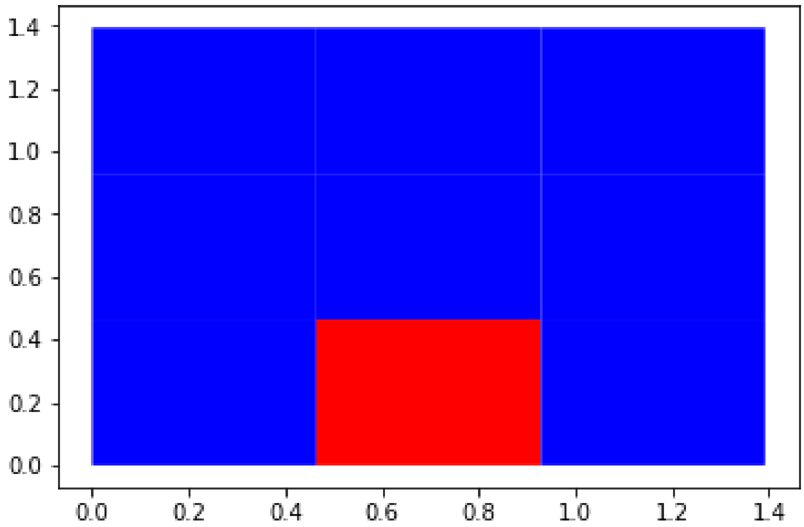


Figura 2: Imagen del espacio clasificado con 10 puntos.

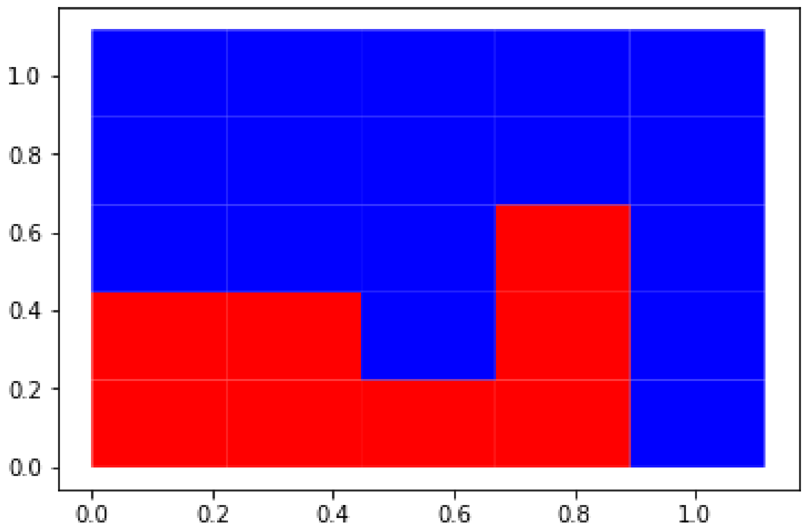


Figura 3: Imagen del espacio clasificado con 100 puntos.

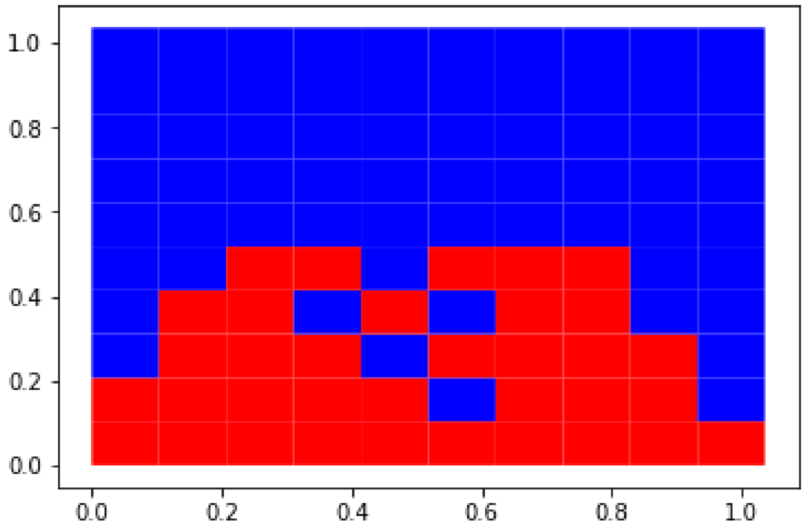


Figura 4: Imagen del espacio clasificado con 1000 puntos.

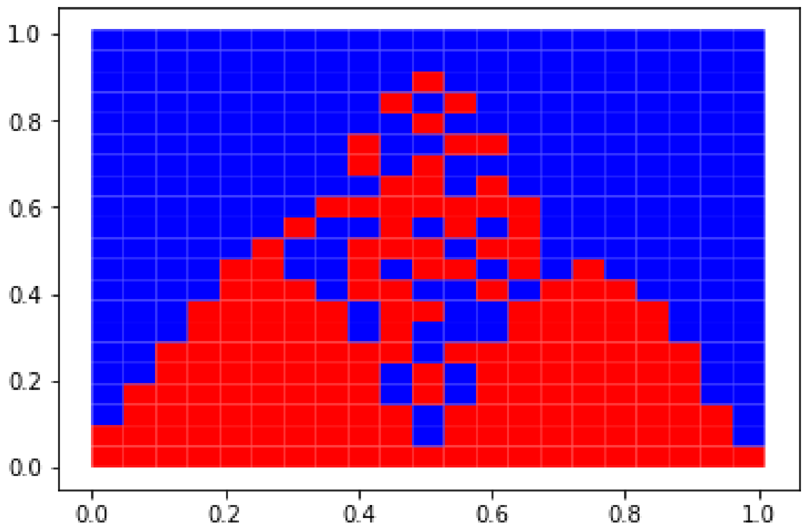


Figura 5: Imagen del espacio clasificado con 10000 puntos.

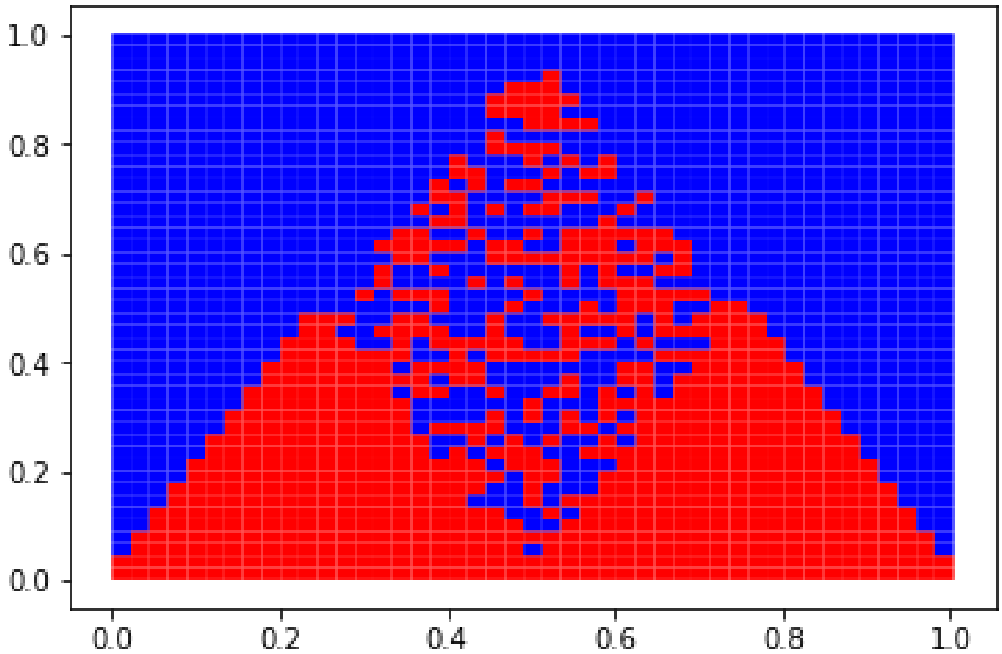


Figura 6: Imagen del espacio clasificado con 100000 puntos.

4. Clase 4

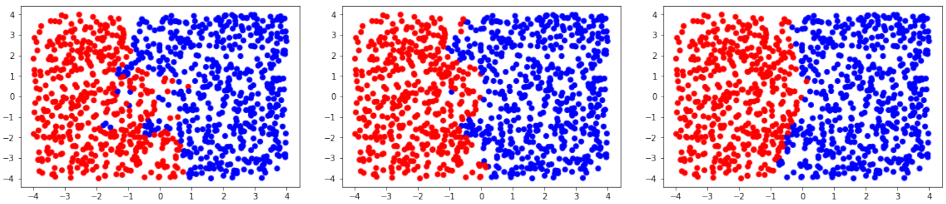
4.1. Ejercicio 9: Seleccion Método de los k-primeros Vecinos

Abajo se simulara distantas cantidades de puntos de color y los cuales se clasificaran con distintos colores de la siguiente forma:

- Se dibujan 50 puntos de color rojo y 50 puntos de puntos de color azul, seguido de N puntos aleatorios los cuales se decidira su color con en base a 100 puntos nombrados.
- Los rojos tendran una distribucion normal multivariada de la forma $\mathcal{N}((-1,0), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$
- Los azules tendran una distribucion normal multivariada de la forma $\mathcal{N}((1,0), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$

- Los puntos aleatorios tienen coordenadas $X, Y \sim \mathcal{U}(-4, 4)$

Ahora el criterio de evaluacion va ser el de los k-primeros vecinos, en el que consiste revisando una cantidad k de vecinos, elegir de que color es el punto evaluado. El experimento se realizara con k = 1, 3 y 13. La cantidad de puntos N = 100, 1000, 10000, 100000.

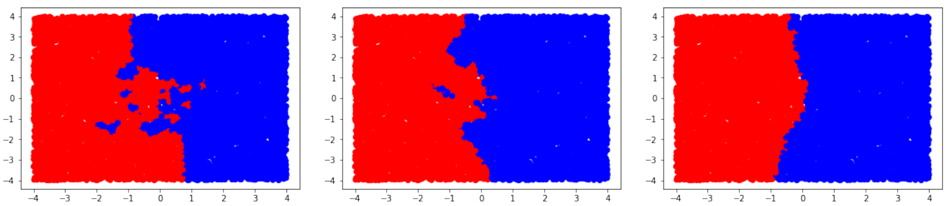


(a) Clasificado usando 1 vecino.

(b) Clasificado usando 3 vecinos.

(c) Clasificado usando 13 vecinos.

Figura 7: Simulacion usando mil puntos

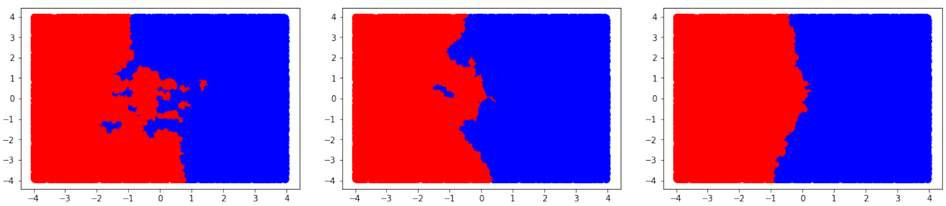


(a) Clasificado usando 1 vecino.

(b) Clasificado usando 3 vecinos.

(c) Clasificado usando 13 vecinos.

Figura 8: Simulacion usando diez mil puntos



(a) Clasificado usando 1 vecino.

(b) Clasificado usando 3 vecinos.

(c) Clasificado usando 13 vecinos.

Figura 9: Simulacion usando cien mil puntos

4.2. Ejercicio 10

Desmostrar:

$$(a+b+c)^2 \leq 3(a^2+b^2+c^2)\forall a,b,c \in \mathbb{R}$$

Desmostracion:

Distribuyendo el cuadrado en el trinomio resulta en lo siguiente:

$$a^2+2ab+2ac+b^2+2bc+c^2 \leq 3(a^2+b^2+c^2)$$

$$(a^2+b^2+c^2)+2ab+2ac+2bc \leq 3(a^2+b^2+c^2)$$

$$2ab+2ac+2bc \leq 2(a^2+b^2+c^2)$$

Para simplificar las cuentas, reescribo a y c como $a = b+\alpha$ y $c = b+\beta$ $\forall \alpha, \beta \in \mathbb{R}$. Quedando la entonces como:

$$b^2+\alpha b+b^2+\beta b+\alpha b+\alpha\beta+b^2+\beta b \leq b^2+2\alpha b+\alpha^2+b^2+b^2+2\beta b+\beta^2$$

resultando en:

$$\alpha\beta \leq \alpha^2+\beta^2$$

Lo cual es siempre verdadero $\forall \alpha, \beta \in \mathbb{R}$.

5. Clase 5

5.1. Ejercicio 11

Dada las desigualdades:

$$P[S_n - E[S_n] \leq \epsilon] \geq e^{-2\epsilon/\sum_{i=1}^{n}(b_i-a_i)^2}$$

$$P[S_n - E[S_n] \geq -\epsilon] \geq e^{-2\epsilon/\sum_{i=1}^{n}(b_i-a_i)^2}$$

Probar que es obvio lo siguiente:

$$P[|S_n - E[S_n]| > \epsilon] \geq 2e^{-2\epsilon/\sum_{i=1}^{n}(b_i-a_i)^2}$$

Solucion, usando la propiedad $P(A \cup B) \leq P(A)+P(B)$ puedo escribir las dos ecuaciones de la siguiente forma:

$$P[S_n - E[S_n] \leq \epsilon \cup S_n - E[S_n] \geq -\epsilon] \leq P[S_n - E[S_n] \leq \epsilon] + P[S_n - E[S_n] \geq -\epsilon]$$

$$\leq e^{-2\epsilon/\sum\limits_{i=1}^{n}(b_i-a_i)^2} + P[S_n - E[S_n] \geq -\epsilon]$$

$$\leq 2e^{-2\epsilon/\sum\limits_{i=1}^{n}(b_i-a_i)^2}$$

Y que inicialmente tengo la $P[-\epsilon \leq S_n - E[S_n] \leq \epsilon]$ lo tanto:

$$P[|S_n - E[S_n]| > \epsilon] \geq 2e^{-2\epsilon/\sum\limits_{i=1}^{n}(b_i-a_i)^2}$$

5.2. Ejercicio 12

Probar la siguiente desigualdad:

$$|\widehat{L}_n(\phi_n^*) - L(\phi_n^*)| \leq \sup_{\phi \in \mathscr{C}} |\widehat{L}_n(\phi) - L(\phi)|$$

Basicamente dice que la que la diferencia entre el minimo error estimado, y el verdadero (basado en los datos), es menor o igual la maxima diferencia que puede haber entre el error estimado y el verdadero para un clasificador determinado.

5.3. Ejercicio 14: Simulacion de error

6. Clase 6

6.1. Ejercicio 16

Para un conjunto $A' = \{(a,b) : a < b\}$, calcular $N_{A'}(z_1, z_2, z_3)$

Los conjuntos que se pueden hacer son $\{\emptyset\}$, $\{z_1\}$, $\{z_2\}$, $\{z_3\}$, $\{z_1, z_2\}$, $\{z_2, z_3\}$, $\{z_1, z_2, z_3\}$. Por lo tanto $N_{A'}(z_1, z_2, z_3) = 7$