Notación

- MSE: del inglés, *Mean Square Error* (error cuadrático medio, en español)
- $\widehat{X}_{\mathrm{MMSE}}$: estimador de mínimo error cuadrático medio
- $\widehat{X}_{\mathrm{MAD}}$: estimador de mínimo error absoluto medio
- $\widehat{S}_{\mathrm{ML}}$: estimador de máxima verosimilitud
- $\widehat{S}_{\mathrm{MAP}}$: estimador de máximo *a posteriori*

1. Se realiza un test de habilidades matemáticas entre los estudiantes universitarios de España matriculados en un Grado de Ingeniería. Suponga que la puntuación X que obtiene un estudiante en la prueba de "estadística y probabilidad" es un número en el intervalo $[0, 1]$, mientras que la puntuación Y que obtiene en la prueba de "funciones y gráficas" también es un número en el intervalo $[0, 1]$.

   Asuma que las puntuaciones X e Y se distribuyen conforme a la siguiente función de densidad probabilidad conjunta:

$$p_{X,Y}(x,y) = \begin{cases} 12\,(1-x)\,y\,(1-y) & 0 \leq x \leq 1 \quad 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

   (a) Establezca las expresiones de las funciones de densidad de probabilidad marginales: $p_X(x)$ y $p_Y(y)$.

   (b) Obtenga el estimador MMSE de X dado Y, $\widehat{X}_{\mathrm{MMSE}}(Y)$.

   (c) Determine el estimador MAD de X dado Y, $\widehat{X}_{\mathrm{MAD}}(Y)$.

   (d) Calcule el MSE de los estimadores obtenidos en los apartados (b) y (c). ¿Cuál de los dos estimadores tiene menor MSE? Justifique su respuesta.

   (e) Si se seleccionase un estudiante universitario al azar, ¿cuál sería el valor predicho de su puntuación en la prueba de "estadística y probabilidad" que tiene menor MSE? Razone su respuesta.

Solution:

(a) $p_X(x) = 2(1-x) \qquad 0 \leq x \leq 1$

$p_Y(y) = 6\,y\,(1-y) \qquad 0 \leq y \leq 1$

(b) $\widehat{X}_{\mathrm{MMSE}} = \dfrac{1}{3}$

(c) $\widehat{X}_{\mathrm{MAD}} = 1 - \dfrac{\sqrt{2}}{2} \approx 0{,}2928$

(d) $\mathbb{E}\left\{\left(X - \widehat{X}_{\mathrm{MMSE}}\right)^2\right\} = \dfrac{1}{18} \approx 0{,}0555$

$\mathbb{E}\left\{\left(X - \widehat{X}_{\mathrm{MAD}}\right)^2\right\} = \dfrac{1}{18} + (0, 0404)^2 \approx 0{,}0571$

El estimador con menor MSE es el estimador $\widehat{X}_{\mathrm{MMSE}}$.

(e) $\widehat{X}_{\mathrm{MMSE}} = \dfrac{1}{3}$ (es el estimador MMSE de X).

(Puntuación: 1,5/4 puntos)

2. Las variables aleatorias X y S están relacionadas a través de la siguiente función de verosimilitud:

$$p_{X|S}(x|s) = \ln(s)s^{-x}, \qquad x \geq 0, \qquad s > 1$$

y la distribución *a priori*:

$$p_S(s) = \frac{1}{s^2}, \qquad s > 1$$

(a) Determine el estimador ML de S a la vista de X, $\widehat{S}_{\mathrm{ML}}(X)$.

(b) Determine el estimador ML de S a la vista de un conjunto $\mathcal{C} = \left\{x^{(k)}\right\}_{k=0}^{K-1}$ de K realizaciones independientes de X.

(c) Determine el estimador MAP de S a la vista de X, $\widehat{S}_{\mathrm{MAP}}(X)$.

Solution:

(a)

$$\begin{aligned}\widehat{S}_{\mathrm{ML}} &= \arg\max_s \ln p_{X|S}(x|s) = \arg\max_s \left(\ln(\ln s) - x\ln(s)\right)\\ &= \exp\left(\frac{1}{x}\right)\end{aligned}$$

(b)

$$\begin{aligned}\widehat{S}_{\mathrm{ML}} &= \arg\max_s \sum_{k=0}^{K-1} \ln p_{X|S}(x^{(k)}|s) = \arg\max_s \left(\sum_{k=0}^{K-1}\left(\ln(\ln s) - x^{(k)}\ln(s)\right)\right)\\ &= K\ln(\ln s) - \ln(s)\sum_{k=0}^{K-1} x^{(k)}\\ &= \exp\left(\frac{K}{\sum_{k=0}^{K-1} x^{(k)}}\right)\end{aligned}$$

(c)

$$\begin{aligned}\widehat{S}_{\mathrm{MAP}} &= \arg\max_s p_{X|S}(x|s)p_S(s) = \arg\max_s \left(\ln(\ln s) - (x+2)\ln(s)\right)\\ &= \exp\left(\frac{1}{x+2}\right)\end{aligned}$$

(Puntuación: 0,75/4 puntos)

3. Considere un problema de decisión binaria con hipótesis equiprobables definido mediante las siguientes funciones de verosimilitud:

$$p_{X_1,X_2|H}(x_1,x_2|H=0) = \begin{cases} 18, & 0 < x_1 < \frac{1}{3}, \quad 0 < x_2 < \frac{1}{3} - x_1 \\ 0, & \text{resto} \end{cases}$$

$$p_{X_1,X_2|H}(x_1,x_2|H=1) = \begin{cases} 6\,(1 - x_1 - x_2), & 0 < x_1 < 1, \quad 0 < x_2 < 1 - x_1 \\ 0, & \text{resto} \end{cases}$$

y política de costes donde acertar no conlleva penalización alguna y $c_{01} = 4c_{10}$.

(a) Determine el decisor de mínimo coste medio dada la observación.

(b) ¿Cuánto deberían valer las probabilidades *a priori* de cada hipótesis para que el decisor de coste medio mínimo, dada la observación, no incurra en falsas alarmas?

Solution:

(a)

$$\begin{cases} D = 1, & 0 \leq x_1 + x_2 \leq \frac{1}{4} \\ D = 0, & \frac{1}{4} \leq x_1 + x_2 \leq \frac{1}{3} \\ D = 1, & \frac{1}{3} \leq x_1 + x_2 \leq 1 \end{cases}$$

(b) $P(H = 0) \geq \frac{4}{7}$

(Puntuación: 0,75/4 puntos)

4. Describa brevemente en qué consiste la validación cruzada e indique para qué se puede emplear.

Solution:

(Puntuación: 0,3/4 puntos)

5. Considere el problema de clasificación definido por el conjunto de entrenamiento que se muestra en la tabla a continuación.

x	Clase
-0,45	-1
-0,03	-1
-0,09	1
0,23	1
0,38	-1
0,92	1
1,19	1
1,28	1
1,33	1

(a) Dibuje la frontera de clasificación que se obtendría con un árbol de decisión de 3 nodos hoja.

(b) Calcule la probabilidad de falsa alarma del clasificador del apartado anterior para el conjunto de test de la tabla siguiente.

x	Clase
-0,15	-1
0,31	-1
1,22	1

Solution:

(Puntuación: 0,4/4 puntos)

6. Describa cómo puede resolverse un problema de predicción de valores de una señal utilizando un filtro lineal. Dibuje un esquema del filtro indicando claramente las señales involucradas.

Solution:

(Puntuación: 0,3/4 puntos)