Cuaderno de trabajo:

Razonamiento probabilístico

Albert Sanchis

Departamento de Sistemas
Informáticos y Computación

Objetivos formativos

- Inferir conocimiento probabilístico mediante las reglas suma y producto del cálculo de probabilidades
- Inferir conocimiento a partir de variables continuas
- Aplicar la regla de decisión de Bayes
- Calcular la probabilidad de error
- Inferir conocimiento probabilístico con el teorema de Bayes

- ***Cuestión 1***: Basándote en la tabla de probabilidades conjuntas del ejemplo del dentista que se muestra a la derecha, y aplicando la regla suma o la regla producto, calcula las siguientes probabilidades:

d	c	h	P
0	0	0	0,576
0	0	1	0,008
0	1	0	0,144
0	1	1	0,072
1	0	0	0,064
1	0	1	0,012
1	1	0	0,016
1	1	1	0,108

1. Probabilidad de observar caries y dolor (a la vez):

   $$P(c=1, d=1) = \sum_{h=0,1} P(h, c=1, d=1) = 0{,}124$$

2. Probabilidad de observar dolor:

   $$P(d=1) = \sum_{h=0,1} \sum_{c=0,1} P(h, c, d=1) = 0{,}2$$

3. Probabilidad de observar caries tras observar (sabiendo que hay) dolor:

   $$P(c=1 \mid d=1) = \frac{P(c{=}1, d{=}1)}{P(d{=}1)} = \frac{0{,}124}{0{,}2} = 0{,}62$$

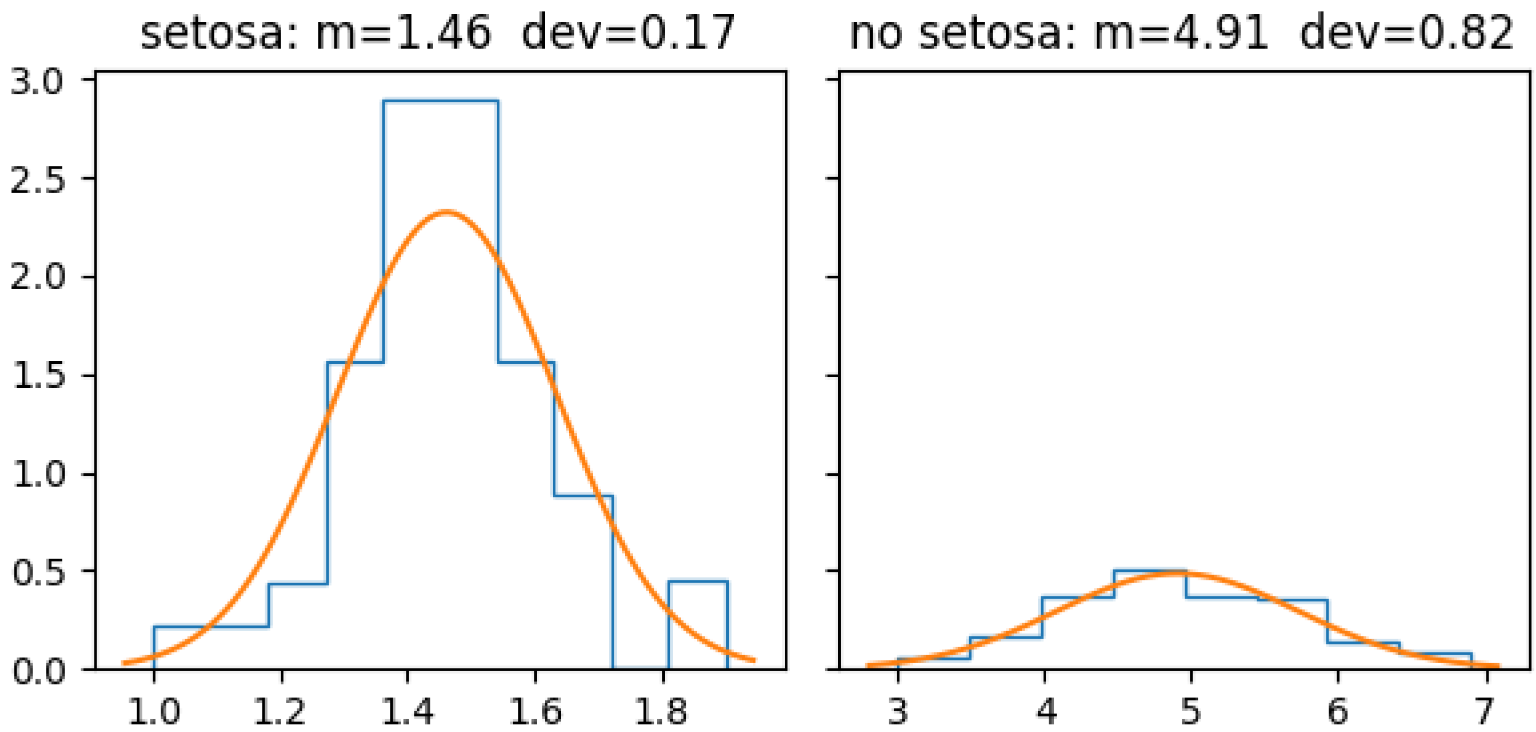
4. Probabilidad de no observar hueco tras observar (sabiendo que hay) dolor:

   $$P(h=0 \mid d=1) = \frac{P(h{=}0, d{=}1)}{P(d{=}1)} = \frac{0{,}08}{0{,}2} = 0{,}4$$

- ***Cuestión 2***: Considera el problema de clasificar flores iris en setosa o no-setosa a partir de su longitud de pétalos (x). El estudio empírico siguiente muestra que las distribuciones de x para setosas y no-setosas pueden aproximarse con distribuciones normales de medias y desviaciones estándares:

$$p(x \mid c = \text{set}) \sim \mathcal{N}(\mu_{\text{set}} = 1{,}46, \sigma_{\text{set}} = 0{,}17)$$

$$p(x \mid c = \text{nos}) \sim \mathcal{N}(\mu_{\text{nos}} = 4{,}91, \sigma_{\text{nos}} = 0{,}82)$$



Asumiendo que las densidades normales estimadas son ciertas y la probabilidad a priori de setosa es $1/3$, contesta a las siguientes preguntas:

1. ¿Cuál es la probabilidad a posteriori de que una flor de longitud de pétalos 2 sea setosa sabiendo que $\mathcal{N}(x = 2 \mid \mu_{\text{set}} = 1{,}46, \sigma_{\text{set}} = 0{,}17) = 0{,}015117$ y $\mathcal{N}(x = 2 \mid \mu_{\text{nos}} = 4{,}91, \sigma_{\text{nos}} = 0{,}82) = 0{,}00089614$?

$$P(c = \textbf{\textit{set}} \mid x = 2) = \frac{P(c = \textbf{\textit{set}})\, p(x = 2 \mid c = \textbf{\textit{set}})}{p(x = 2)}$$

$$= \frac{P(c = \textbf{\textit{set}})\, p(x = 2 \mid c = \textbf{\textit{set}})}{P(c = \textbf{\textit{set}})\, p(x = 2 \mid c = \textbf{\textit{set}}) + P(c = \textbf{\textit{nos}})\, p(x = 2 \mid c = \textbf{\textit{nos}})}$$

$$= \frac{1/3 \cdot \mathcal{N}(x = 2 \mid \mu_{\textbf{\textit{set}}} = 1{,}46, \sigma_{\textbf{\textit{set}}} = 0{,}17)}{1/3 \cdot \mathcal{N}(x = 2 \mid \mu_{\textbf{\textit{set}}} = 1{,}46, \sigma_{\textbf{\textit{set}}} = 0{,}17) + 2/3 \cdot \mathcal{N}(x = 2 \mid \mu_{\textbf{\textit{nos}}} = 4{,}91, \sigma_{\textbf{\textit{nos}}} = 0{,}82)}$$

$$= \frac{1/3 \cdot 0{,}015117}{1/3 \cdot 0{,}015117 + 2/3 \cdot 0{,}00089614} = 0{,}89$$

2. ¿Cuál es la decisión óptima de clasificación de esta flor?

$$c^*(x=2) = \arg\max_c \begin{pmatrix} P(c = \textbf{\textit{set}} \mid x=2) & = 0{,}89 \\ P(c = \textbf{\textit{nos}} \mid x=2) & = 0{,}11 \end{pmatrix} = \textbf{\textit{set}}$$

3. ¿Cuál es la probabilidad de qué dicha decisión sea errónea?

$$\begin{aligned} P(\textbf{\textit{error}} \mid x=2) &= 1 - P(c^*(x=2) \mid x) \\ &= 1 - P(c = \textbf{\textit{set}} \mid x) \\ &= 1 - 0{,}89 \\ &= 0{,}11 \end{aligned}$$

- ***Cuestión 3***: Teniendo en cuenta la siguiente información sobre la enfermedad de la meningitis:
  - La meningitis causa rigidez de nuca en un $70\,\%$ de los casos.
  - La probabilidad a priori de que un paciente tenga meningitis es de $1\,/\,100\,000$.
  - La probabilidad a priori de que un paciente tenga rigidez de nuca es del $1\,\%$.

  Calcula la probabilidad de que un paciente con rigidez de nuca tenga meningitis.

$$P(m=1 \mid r=1) = \frac{P(m=1)\cdot P(r=1 \mid m=1)}{P(r=1)}$$

$$= \frac{1/100000 \cdot 70/100}{1/100}$$

$$= \frac{7}{10000} = 0{,}0007$$