T6 EJERCICIOS

Neus Molero

24 de enero de 2021

Ejercicio 1

La red bayesiana de la figura 1 representa de forma muy simplificada el problema del diagnóstico del cáncer de pulmón (obtenido de *Bayesian Artificial Intelligence*. Kevin B. Korb and Ann E. Nicholson, CRC Press, 2010),

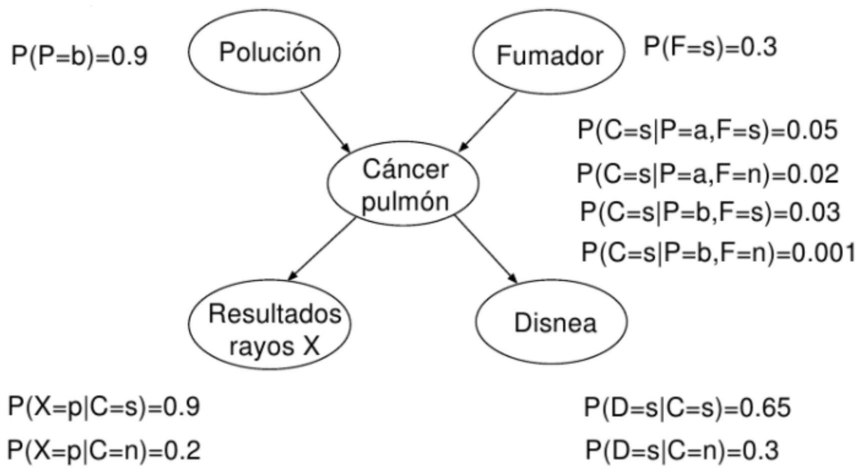

Figura 1: Red bayesiana del ejercicio 1

donde "Polución" puede tomar los valores "b" (bajo) o "a" (alto), "Fumador" puede tomar los valores "s" (sí) o "n" (no), "Resultados de los rayos X" puede tomar los valores "p" (positivo) o "n" (negativo), "Disnea" puede tomar los valores "s" (sí) o "n" (no) y "Cáncer de pulmón" puede tomar los valores "s" (sí) o "n" (no). Calcular:

1. La probabilidad de que el paciente sea fumador sabiendo que padece disnea y que los resultados de rayos X han salido negativos;

2. La probabilidad de que un paciente sufra disnea sabiendo que es fumador y que los resultados de rayos X han salido positivos;

3. La probabilidad de que un paciente sufra cáncer y padezca disnea sabiendo que es fumador, el ambiente en el que vive el paciente presenta una polución alta y que los resultados de rayos X han salido positivos.

Solución del ejercicio 1

1. La probabilidad de que el paciente sea fumador sabiendo que padece disnea y que los resultados de rayos X han salido negativos se puede calcular de la manera siguiente:

$$
\begin{aligned}
P(F|D,X) &= \frac{P(F,D,X)}{P(D,X)} \\
&= \frac{\sum_p \sum_c P(P)\ P(F)\ P(C|P,F)\ P(X|C)\ P(D|C)}{\sum_p \sum_f \sum_c P(P)\ P(F)\ P(C|P,F)\ P(X|C)\ P(D|C)} \\
&= \frac{P(F)\ \sum_p P(P)\ \sum_c P(C|P,F)\ P(X|C)\ P(D|C)}{\sum_p P(P)\ \sum_f P(F)\ \sum_c P(C|P,F)\ P(X|C)\ P(D|C)}
\end{aligned}
$$

   En nuestro caso podemos escribir:

$$P(F=s|D=s,X=n)=$$

$$= \frac{P(F=s)\ \sum_p (P=p)\ \sum_c P(C=c|P=p,F=s)\ P(X=n|C=c)\ P(D=s|C=c)}{\sum_p P(P=p)\ \sum_f P(F=f)\ \sum_c P(C=c|P=p,F=f)\ P(X=n|C=c)\ P(D=s|C=c)}$$

   Ahora substituimos los términos de la ecuación anterior por sus valores numéricos:

$$P(F=s|D=s,X=n)=$$

$$= \frac{0{,}3 \cdot (0{,}9 \cdot (0{,}03 \cdot 0{,}1 \cdot 0{,}65 + 0{,}97 \cdot 0{,}8 \cdot 0{,}3) + 0{,}1 \cdot (0{,}05 \cdot 0{,}1 \cdot 0{,}65 + 0{,}95 \cdot 0{,}8 \cdot 0{,}3))}{0{,}9 \cdot (0{,}1 \cdot 0{,}65 \cdot 0{,}0097 + 0{,}8 \cdot 0{,}3 \cdot 0{,}9903) + 0{,}1 \cdot (0{,}1 \cdot 0{,}65 \cdot 0{,}029 + 0{,}8 \cdot 0{,}3 \cdot 0{,}977)}$$

$$= \frac{0{,}0703}{0{,}2381} = 0{,}2953$$

   Por tanto, la probabilidad de que el paciente sea fumador sabiendo que padece disnea y que los resultados de rayos X han salido negativos es, en porcentaje, del 29,53 %

2. La probabilidad de que un paciente sufra disnea sabiendo que es fumador y que los resultados de rayos X han salido positivos la calcularemos de la siguiente forma:

$$
\begin{aligned}
P(D|F,X) &= \frac{P(D,F,X)}{P(F,X)} \\
&= \frac{\sum_p \sum_c P(P)\ \cancel{P(F)}\ P(C|P,F)\ P(X|C)\ P(D|C)}{\sum_p \sum_c \sum_d P(P)\ \cancel{P(F)}\ P(C|P,F)\ P(X|C)\ P(D|C)} \\
&= \frac{\sum_p P(P)\ \sum_c P(C|P,F)\ P(X|C)\ P(D|C)}{\sum_p P(P)\ \sum_c P(C|P,F)\ P(X|C)\ \cancelto{1}{\sum_d P(D|C)}} \\
&= \frac{\sum_p P(P)\ \sum_c P(C|P,F)\ P(X|C)\ P(D|C)}{\sum_p P(P)\ \sum_c P(C|P,F)\ P(X|C)}
\end{aligned}
$$

   Podemos particularizar la expresión anterior y adaptarla a nuestro problema:

$$P(D=s|F=s,X=p)=$$

$$= \frac{\sum_p P(P=p)\ \sum_c P(C=c|P=p,F=s)\ P(X=p|C=c)\ P(D=s|C=c)}{\sum_p P(P=p)\ \sum_c P(C=c|P=p,F=s)\ P(X=p|C=c)}$$

Ahora substituimos los términos de la ecuación anterior por sus valores numéricos:

$$P(D = s|F = s, X = p) =$$

$$= \frac{0{,}9 \cdot (0{,}03 \cdot 0{,}65 \cdot 0{,}9 + 0{,}97 \cdot 0{,}3 \cdot 0{,}2) + 0{,}1 \cdot (0{,}05 \cdot 0{,}65 \cdot 0{,}9 + 0{,}95 \cdot 0{,}3 \cdot 0{,}2)}{0{,}9 \cdot (0{,}03 \cdot 0{,}9 + 0{,}97 \cdot 0{,}2) + 0{,}1 \cdot (0{,}05 \cdot 0{,}9 + 0{,}095 \cdot 0{,}2)}$$

$$= \frac{0{,}0768}{0{,}2224} = 0{,}3453$$

Por consiguiente, la probabilidad de que un paciente sufra disnea sabiendo que es fumador y que los resultados de rayos X han salido positivos resulta, en porcentaje, 34,53 %

3. Para calcular la probabilidad de que un paciente sufra cáncer y padezca disnea sabiendo que es fumador, el ambiente en el que vive el paciente presenta una polución alta y que los resultados de rayos X han salido positivos, partimos de la expresión $P(C \cap D|F, P, X)$.

   Teniendo en cuenta la expresión matemática del cálculo de la probabilidad condicionada $P(A|B) = P(A \cap B)/P(B)$, podemos escribir para nuestro caso:

$$\begin{aligned}
P(C, D|F, P, X) &= \frac{P(C, D, F, P, X)}{P(F, P, X)} \\
&= \frac{P(P)\ P(F)\ P(C|P, F)\ P(X|C)\ P(D|C)}{\sum_c \sum_d P(P)\ P(F)\ P(C|P, F)\ P(X|C)\ P(D|C)} \\
&= \frac{\cancel{P(P)}\ \cancel{P(F)}\ P(C|P, F)\ P(X|C)\ P(D|C)}{\sum_c \cancel{P(P)}\ \cancel{P(F)}\ P(C|P, F)\ P(X|C)\ \sum_d \cancel{P(D|C)}^{\,1}} \\
&= \frac{P(C|P, F)\ P(X|C)\ P(D|C)}{\sum_c P(C|P, F)\ P(X|C)}
\end{aligned}$$

A partir de aquí podemos considerar la probabilidad que queremos calcular:

$$P(C = s, D = s|F = s, P = a, X = p) =$$

$$= \frac{P(C = s|P = a, F = s)\ P(X = p|C = s)\ P(D = s|C = s)}{\sum_c P(C = c|P = a, F = s)\ P(X = p|C = c)}$$

Ahora substituimos los términos de la ecuación anterior por sus valores numéricos:

$$P(C = s, D = s|F = s, P = a, X = p) =$$

$$= \frac{0{,}05 \cdot 0{,}9 \cdot 0{,}65}{(0{,}05 \cdot 0{,}9) + (0{,}95 \cdot 0{,}2)}$$

$$= \frac{0{,}02925}{0{,}235} = 0{,}1245$$

Así pues, la probabilidad de que un paciente sufra cáncer y padezca disnea sabiendo que es fumador, el ambiente en el que vive el paciente presenta una polución alta y que los resultados de rayos X han salido positivos es del 12,45 %.