

Modelos gráficos probabilísticos

Marcos Esteve Casademunt

Diciembre 2018

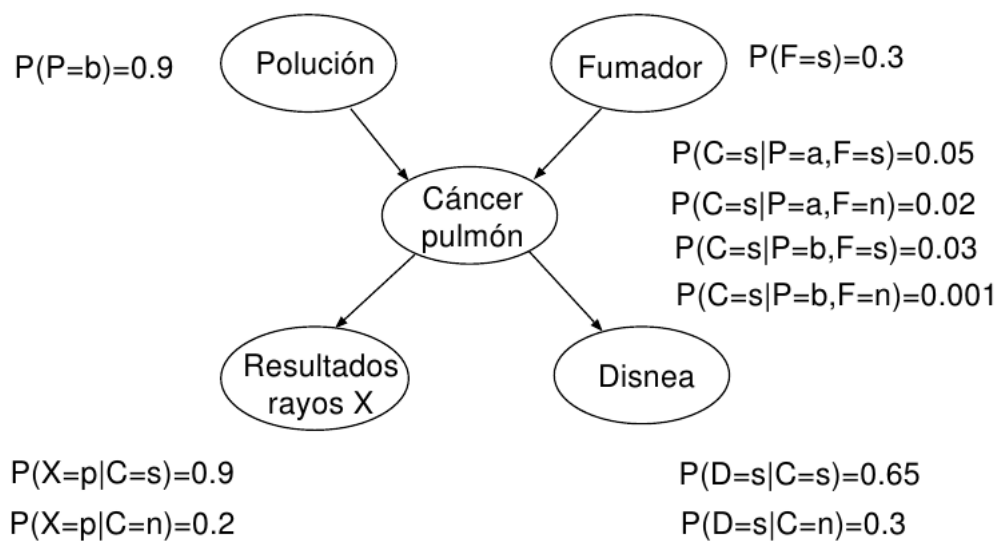


Figura 1: Red Bayesiana para el diagnóstico de cáncer de pulmón

La distribución de probabilidad conjunta asociada a la siguiente Red Bayesiana es:

$$P(P, F, C, X, D) = P(P) * P(F) * P(C|P, F) * P(X|C) * P(D|C) \quad (1)$$

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).

1. Ejercicio 1

Se nos pide calcular la probabilidad de que el paciente sea fumador sabiendo que padece disnea y que los resultados de rayos X han salido negativos. Es decir

$$\begin{aligned}
P(F = s|D = s, X = n) &= \frac{P(F = s, D = s, X = n)}{P(D = s, X = n)} = \\
&= \frac{\sum_{\forall p,c} P(P = p, F = s, C = c, X = n, D = s)}{\sum_{\forall p,c,f} P(P = p, F = f, C = c, X = n, D = s)} = \\
&= \frac{\sum_{\forall p,c} P(P = p) * P(F = s) * P(C = c|P = p, F = s) * P(X = n|C = c) * P(D = s|C = c)}{\sum_{\forall p,c,f} P(P = p) * P(F = f) * P(C = c|P = p, F = f) * P(X = n|C = c) * P(D = s|C = c)} = \\
&= \frac{,9 * ,3 * ,03 * ,1 * ,65 + ,9 * ,3 * ,97 * ,8 * ,3 + ,1 * ,3 * ,05 * ,1 * ,65 + ,1 * ,3 * ,95 * ,8 * ,3}{,07032 + ,9 * ,7 * ,01 * ,1 * ,65 + ,9 * ,7 * ,99 * ,8 * ,3 + ,1 * ,7 * ,02 * ,1 * ,65 + ,1 * ,7 * ,98 * ,8 * ,3} = \\
&= \frac{0,07032}{0,2369725} = 0,2967 \quad (2)
\end{aligned}$$

Por tanto, la probabilidad de que el paciente sea fumador sabiendo que tiene disnea y los rayos X han salido negativo es de **0,2967**

2. Ejercicio 2

Se nos pide calcular la probabilidad de que un paciente sufra disnea sabiendo que es fumador y que los resultados de rayos X han salido positivos. Es decir:

$$\begin{aligned}
P(D = s|F = s, X = s) &= \frac{P(F = s, D = s, X = s)}{P(F = s, X = s)} = \\
&= \frac{\sum_{\forall p,c} P(P = p, F = s, C = c, X = s, D = s)}{\sum_{\forall p,c,d} P(P = p, F = s, C = c, X = s, D = d)} = \\
&= \frac{\sum_{\forall p,c} P(P = p) * P(F = s) * P(C = c|P = p, F = s) * P(X = s|C = c) * P(D = s|C = c)}{\sum_{\forall p,c,d} P(P = p) * P(F = s) * P(C = c|P = p, F = w) * P(X = s|C = c) * P(D = d|C = c)} = \\
&= \frac{,9 * ,3 * ,03 * ,65 * ,9 + ,9 * ,3 * ,97 * ,3 * ,2 + ,1 * ,3 * ,05 * ,65 * ,9 + ,1 * ,3 * ,95 * ,3 * ,2}{,9 * ,3 * ,03 * ,35 * ,9 + ,9 * ,3 * ,97 * ,7 * ,2 + ,1 * ,3 * ,02 * ,35 * ,9 + ,1 * ,3 * ,98 * ,7 * ,2} = \\
&= \frac{0,02304}{0,0665} = 0,3464 \quad (3)
\end{aligned}$$

La probabilidad de que el paciente sufra disnea sabiendo que es fumador y que los rayos X han salido positivos es de **0,3464**

3. Ejercicio 3

Se pide calcular la probabilidad de que un paciente sufra cáncer sabiendo que es fumador, sufre disnea y que los resultados de rayos X han salido

positivos. Es decir:

$$\begin{aligned}
 P(C = s|F = s, X = s, D = s) &= \frac{P(C = s, F = s, X = s, D = s)}{P(F = s, X = s, D = s)} = \\
 &= \frac{\sum_{\forall p} P(P = p, F = s, C = s, X = s, D = s)}{\sum_{\forall p, c} P(P = p, F = s, C = c, X = s, D = s)} = \\
 &= \frac{\sum_{\forall p} P(P = p) * P(F = s) * P(C = s|P = p, F = s) * P(X = s|C = s) * P(D = s|C = s)}{\sum_{\forall p, c} P(P = p) * P(F = s) * P(C = c|P = p, F = s) * P(X = s|C = c) * P(D = s|C = c)} = \\
 &= \frac{,9 * ,3 * ,03 * ,65 * ,9 + ,1 * ,3 * ,05 * ,65 * ,09}{0,02304} = 0,24375 \quad (4)
 \end{aligned}$$

La probabilidad asociada al siguiente ejercicio será por tanto: **0.24375**