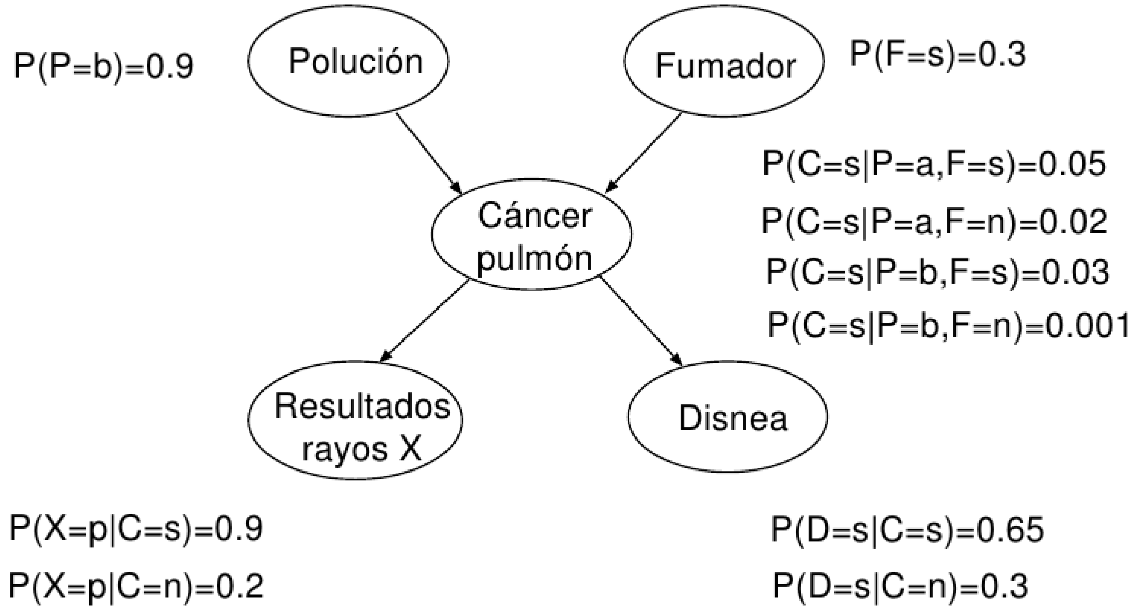
TRABAJO TEMA 6 APR

EJERCICIO 1.- La red bayesiana de la figura 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),



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.

APARTADO 1

Primero obtenemos la expresión simplificada:

$$P(F|D,X) = \frac{P(D,F,X)}{P(D,X)}$$

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

En segundo lugar, sustituimos por aquellos valores conocidos:

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

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

Donde para simplificar diremos que:

$$i = P(F = s) \sum\nolimits_p P(P = p) \sum\nolimits_c P(C = c|P = p, F = s) P(X = n|C = c) P(D = s|C = c)$$

$$j = P(F = n) \sum\nolimits_p P(P = p) \sum\nolimits_c P(C = c|P = p, F = n) P(X = n|C = c) P(D = s|C = c)$$

Dejando la siguiente expresión:

$$P(F = s|D = s, X = n) = \frac{i}{i + j}$$

Finalmente, sustituimos los términos por los valores de las probabilidades:

$i = 0.3\big(0.9(0.03 * 0.1 * 0.65 + 0.97 * 0.8 * 0.3) + 0.1(0.05 * 0.1 * 0.65 + 0.95 * 0.8 * 0.3)\big)$
$i = 0.07032$

$j = 0.7\big(0.9(0.001 * 0.1 * 0.65 + 0.999 * 0.8 * 0.3) + 0.1(0.02 * 0.1 * 0.65 + 0.98 * 0.8 * 0.3)\big)$
$j = 0.167645$

$$(F = s|D = s, X = n) = \frac{0.07032}{0.07032 + 0.167645} = 0.295506$$

Por ende, la solución dada en porcentaje sería del 29.5506%.

APARTADO 2

Primero obtenemos la expresión simplificada:

$$P(D|F, X) = \frac{P(D, F, X)}{P(F, X)}$$

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

$$= \frac{\sum_p P(P = p) \sum_c P(C = c|P = p, F) P(X|C = c) P(D|C = c)}{\sum_p P(P = p) \sum_c P(C = c|P = p, F) P(X|C = c) * 1}$$

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

En segundo lugar, sustituimos por aquellos valores conocidos:

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

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

Finalmente, sustituimos los términos por los valores de las probabilidades:

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

$$= \frac{0.9(0.03 * 0.65 * 0.9 + 0.97 * 0.3 * 0.2) + 0.1(0.05 * 0.65 * 0.9 + 0.95 * 0.3 * 0.2)}{0.9(0.03 * 0.9 + 0.97 * 0.2) + 0.1(0.05 * 0.9 + 0.95 * 0.2)}$$

$$= \frac{0.0768}{0.2224} = 0.345324$$

Por ende, la solución dada en porcentaje sería del 34.5324%.

APARTADO 3

Primero obtenemos la expresión simplificada:

$$P(C, D|F, P, X) = \frac{P(C, D, F, P, X)}{P(F, P, X)}$$

$$= \frac{\cancel{P(F)P(P)}P(C|P, F)P(X|C)P(D|C)}{\cancel{P(F)P(P)}\sum_c P(C = c|P, F)P(X|C = c)\sum_d P(D = d|C = c)}$$

$$= \frac{P(C|P, F)P(X|C)P(D|C)}{\sum_c P(C = c|P, F)P(X|C = c) * 1}$$

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

En segundo lugar, sustituimos por aquellos valores conocidos:

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

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

Finalmente, sustituimos los términos por los valores de las probabilidades:

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

$$= \frac{0.05 * 0.9 * 0.65}{0.05 * 0.9 + 0.95 * 0.2} = \frac{0.02925}{0.235} = 0.124468$$

Por ende, la solución dada en porcentaje sería del 12.4468%.