Tarea 8: Razonamiento con Probabilidades

Aplicar razonamiento probabilístico para calcular la probabilidad de que una persona tenga Covid de acuerdo al siguiente conocimiento:

- **Factores que influyen (supondremos que son independientes):**
  - **La tasa de infectados en la zona donde se vive. Consideraremos tres tipos de zona, zona de incidencia baja, zona de incidencia media y zona de incidencia alta.**
  - **Si la persona está vacunada o no.**
- **Efectos o evidencias que constatar (supondremos que son independientes):**
  - **Fiebre: consideraremos tres casos, sin fiebre, con fiebre moderada y con fiebre alta.**
  - **Tos: consideraremos dos casos, con tos y sin tos.**

1. **Indicar las distribuciones de probabilidad con las que haría falta contar para realizar el razonamiento. Suponer que se obtienen y actualizan a partir de un banco de datos estadísticos, y asignar unos valores concretos (probabilidad subjetiva) para cada una de esas probabilidades necesarias.**

Introducimos la siguiente notación para una lectura más clara:

- Covid ($\boldsymbol{c}$): sí ($+\boldsymbol{c}$), no ($-\boldsymbol{c}$).
- Factores:
  - Incidencia de la zona donde vive ($\boldsymbol{z}$): alta ($\boldsymbol{z_a}$), media ($\boldsymbol{z_m}$), baja ($\boldsymbol{z_b}$).
  - Vacunada: sí ($+\boldsymbol{v}$), no ($-\boldsymbol{v}$).
- Efectos:
  - Fiebre: alta ($\boldsymbol{f_a}$), moderada ($\boldsymbol{f_m}$), baja ($\boldsymbol{f_b}$).
  - Tos: sí ($+\boldsymbol{t}$), no ($-\boldsymbol{t}$).

Pasamos a asignar unos valores concretos (probabilidad subjetiva) para cada una de las probabilidades necesarias. Recolectando información en internet para asignar unos valores razonables, tendríamos lo siguiente:

- **Probabilidad de la variable Covid:**
  - $\boldsymbol{P(+c) = 0.2}$
  - $\boldsymbol{P(-c) = 0.8}$

- **Distribución de probabilidad de los factores:**
  - **Incidencia de la zona:**
    - $\boldsymbol{P(z_a) = 0.55}$
    - $\boldsymbol{P(z_m) = 0.44}$
    - $\boldsymbol{P(z_b) = 0.01}$
  - **Vacunada:**
    - $\boldsymbol{P(+v) = 0.5}$
    - $\boldsymbol{P(-v) = 0.5}$

- **Distribución de probabilidad condicionada de la variable con respecto a los factores:**
  - $\boldsymbol{P(+c \,/\, z_a, +v) = 0.01}$
  - $\boldsymbol{P(+c \,/\, z_a, -v) = 0.85}$
  - $\boldsymbol{P(+c \,/\, z_m, +v) = 0.005}$
  - $\boldsymbol{P(+c \,/\, z_m, -v) = 0.55}$
  - $\boldsymbol{P(+c \,/\, z_b, +v) = 0.001}$
  - $\boldsymbol{P(+c \,/\, z_b, -v) = 0.2}$

- **Distribución de probabilidad condicionada de los efectos con respecto a la variable:**
  - $\boldsymbol{P(f_a \,/\, +c) = 0.85}$
  - $\boldsymbol{P(f_a \,/\, -c) = 0.005}$
  - $\boldsymbol{P(f_m \,/\, +c) = 0.4}$
  - $\boldsymbol{P(f_m \,/\, -c) = 0.55}$
  - $\boldsymbol{P(f_b \,/\, +c) = 0.1}$
  - $\boldsymbol{P(f_b \,/\, -c) = 0.8}$
  - $\boldsymbol{P(+t \,/\, +c) = 0.99}$
  - $\boldsymbol{P(+t \,/\, -c) = 0.01}$

Con estas distribuciones, ya tendríamos todo lo necesario. Gracias a ellas, podemos deducir las siguientes probabilidades usando las leyes y las propiedades de la probabilidad condicionada:

- $\boldsymbol{P(-t \,/\, +c)} = 1 - P(+t \,/\, +c) =$
  $= 1 - 0.99 = \boldsymbol{0.01}$

- $\boldsymbol{P(-t \,/\, -c)} = 1 - P(+t \,/\, -c) =$
  $= 1 - 0.01 = \boldsymbol{0.99}$

- $\boldsymbol{P(+c, z_a, +v)} = P(z_a) * P(+v) * P(+c \,/\, z_a, +v) =$
  $= 0.55 * 0.5 * 0.01 = \boldsymbol{0.00275}$

- $\boldsymbol{P(+c, z_a, -v)} = P(z_a) * P(-v) * P(+c \,/\, z_a, -v) =$
  $= 0.55 * 0.5 * 0.85 = \boldsymbol{0.23375}$

- $\boldsymbol{P(+c, z_a)} = P(+c, z_a, +v) + P(+c, z_a, -v) =$
  $= 0.00275 + 0.23375 = \mathbf{0.2365}$

- $\boldsymbol{P(+c \;/\; z_a)} = P(+c, z_a) \;/\; P(z_a) =$
  $= 0.2365 \;/\; 0.55 = \mathbf{0.43}$

- $\boldsymbol{P(+c, z_m, +v)} = P(z_m) * P(+v) * P(+c \;/\; z_m, +v) =$
  $= 0.44 * 0.5 * 0.005 = \mathbf{0.0011}$

- $\boldsymbol{P(+c, z_m, -v)} = P(z_m) * P(-v) * P(+c \;/\; z_m, -v) =$
  $= 0.44 * 0.5 * 0.55 = \mathbf{0.121}$

- $\boldsymbol{P(+c, z_m)} = P(+c, z_m, +v) + P(+c, z_m, -v) =$
  $= 0.0011 + 0.121 = \mathbf{0.1221}$

- $\boldsymbol{P(+c \;/\; z_m)} = P(+c, z_m) \;/\; P(z_m) =$
  $= 0.1221 \;/\; 0.44 = \mathbf{0.2775}$

- $\boldsymbol{P(+c, z_b, +v)} = P(z_b) * P(+v) * P(+c \;/\; z_b, +v) =$
  $= 0.01 * 0.5 * 0.001 = \mathbf{0.000005}$

- $\boldsymbol{P(+c, z_b, -v)} = P(z_b) * P(-v) * P(+c \;/\; z_b, -v) =$
  $= 0.01 * 0.5 * 0.2 = \mathbf{0.001}$

- $\boldsymbol{P(+c, z_b)} = P(+c, z_b, +v) + P(+c, z_b, -v) =$
  $= 0.000005 + 0.001 = \mathbf{0.001005}$

- $\boldsymbol{P(+c \;/\; z_b)} = P(+c, z_b) \;/\; P(z_b) =$
  $= 0.001005 \;/\; 0.01 = \mathbf{0.1005}$

- $\boldsymbol{P(+c, +v)} = P(+c, z_a, +v) + P(+c, z_m, +v) + P(+c, z_b, +v) =$
  $= 0.00275 + 0.0011 + 0.000005 = \mathbf{0.003855}$

- $\boldsymbol{P(+c \;/\; +v)} = P(+c, +v) \;/\; P(+v)$
  $= 0.003855 \;/\; 0.5 = \mathbf{0.00771}$

- $\boldsymbol{P(+c, -v)} = P(+c, z_a, -v) + P(+c, z_m, -v) + P(+c, z_b, -v) =$
  $= 0.23375 + 0.121 + 0.001 = \mathbf{0.35575}$

- $\boldsymbol{P(+c \;/\; -v)} = P(+c, -v) \;/\; P(-v)$
  $= 0.35575 \;/\; 0.5 = \mathbf{0.7115}$

2. **Calcular la probabilidad de que una persona tenga Covid si vive en una zona de alta incidencia, no está vacunada y tiene fiebre alta pero no tiene tos.**

Se nos pide calcular $\boldsymbol{P(+c\ /\ z_a, -v, f_a, -t)}$. Por tanto, usando las leyes y las propiedades de la probabilidad condicionada, se tiene que:

- $$P(+c, z_a, -v, f_a, -t) = P(z_a) * P(-v) * P(+c\ /\ z_a, -v)\ * P(f_a\ /\ +c)\ * \\ * P(-t\ /\ +c)\ = \\ = 0.55 * 0.5 * 0.85 * 0.85 * 0.01 = \\ = 0.001986875$$

- $$P(-c, z_a, -v, f_a, -t) = P(z_a) * P(-v) * P(-c\ /\ z_a, -v)\ * P(f_a\ /\ -c)\ * \\ * P(-t\ /\ -c)\ = \\ = 0.55 * 0.5 * (1 - 0.85) * 0.005 * 0.99 = \\ = 0.0002041875$$

- $$P(z_a, -v, f_a, -t) = P(+c, z_a, -v, f_a, -t)\ +\ P(-c, z_a, -v, f_a, -t)\ = \\ = 0.001986875\ + 0.0002041875\ = \\ = 0.0021910625$$

- $$\boldsymbol{P(+c\ /\ z_a, -v, f_a, -t)}\ =\ P(+c, z_a, -v, f_a, -t) \,/\, P(z_a, -v, f_a, -t)\ = \\ = 0.001986875 \,/\, 0.0021910625 = \\ = \boldsymbol{0.9068089112017}$$

3. **Calcular la probabilidad de que una persona no tenga Covid si vive en una zona de alta incidencia, está vacunado, tiene fiebre alta y tiene tos.**

Se nos pide calcular $\boldsymbol{P(-c\ /\ z_a, +v, f_a, +t)}$. Usando las leyes y las propiedades de la probabilidad condicionada, obtenemos lo siguiente:

- $$P(+c, z_a, +v, f_a, +t) = P(z_a) * P(+v) * P(+c\ /\ z_a, +v)\ * P(f_a\ /\ +c)\ * \\ * P(+t\ /\ +c)\ = \\ = 0.55 * 0.5 * 0.01\ * 0.85 * 0.99 = \\ = 0.002314125$$

- $$P(-c, z_a, +v, f_a, +t) = P(z_a) * P(+v) * P(-c\ /\ z_a, +v)\ * P(f_a\ /\ -c)\ * \\ * P(+t\ /\ -c)\ = \\ = 0.55 * 0.5 * (1\ -\ 0.01) * 0.005 * 0.01 = \\ = 0.0000136125$$

- $$P(z_a, +v, f_a, +t) = P(+c, z_a, +v, f_a, +t)\ +\ P(-c, z_a, +v, f_a, +t)\ = \\ = 0.002314125 + 0.0000136125 = \\ = 0.0023277375$$

- $$\boldsymbol{P(-c\ /\ z_a, +v, f_a, +t)}\ =\ P(-c, z_a, +v, f_a, +t) \,/\, P(z_a, +v, f_a, +t)\ = \\ = 0.0000136125 \,/\, 0.0023277375\ = \\ = \boldsymbol{0.0058479532163}$$

4. **Calcular la probabilidad de que una persona tenga Covid si ha llegado desde una zona donde no se conoce la situación, no está vacunada, tiene fiebre alta y tiene tos.**

Se nos pide calcular $\boldsymbol{P(+c\ /\ -v, f_a, +t)}$. Gracias a las leyes y las propiedades de la probabilidad condicionada, deducimos que:

- $P(+c, -v, f_a, +t) = P(-v) * P(+c/-v) * P(f_a/+c) * P(+t/+c) =$
  $= 0.5 * 0.7115 * 0.85 * 0.99 =$
  $= 0.299363625$

- $P(-c, -v, f_a, +t) = P(-v) * P(-c/-v) * P(f_a/-c) * P(+t/-c) =$
  $= 0.5 * (1 - 0.7115) * 0.005 * 0.01 =$
  $= 0.00000072125$

- $P(-v, f_a, +t) = P(+c, -v, f_a, +t) + P(-c, -v, f_a, +t) =$
  $= 0.299363625 + 0.00000072125 =$
  $= 0.29936434625$

- $\boldsymbol{P(+c\ /\ -v, f_a, +t)} = P(+c, -v, f_a, +t)\ /\ P(-v, f_a, +t) =$
  $= 0.299363625\ /\ 0.29936434625 =$
  $= \mathbf{0.9999975907284}$

5. **A la persona anterior se le aplica un test y da negativo. Suponiendo que el resultado del test no se ve afectado por tener fiebre o tos, y sabiendo que P(Test Positivo/Covid) = 0.9 y P(Test Negativo/No Covid) = 0.99, ¿cuál sería ahora la probabilidad de que padezca Covid?**

Tenemos ahora un efecto más a tener en cuenta (independiente con respecto al resto de efectos):

- Test: positivo ($+\boldsymbol{test}$), negativo ($-\boldsymbol{test}$).

Sabemos que:

- $\boldsymbol{P(+test\ /\ +c) = 0.9}$
- $\boldsymbol{P(-test\ /\ -c) = 0.99}$

de donde podemos deducir también lo siguiente:

- $\boldsymbol{P(-test\ /+c)} = 1 - P(+test\ /+c) =$
  $= 1 - 0.9 = \mathbf{0.1}$

- $\boldsymbol{P(+test\ /-c)} = 1 - P(-test\ /-c) =$
  $= 1 - 0.99 = \mathbf{0.01}$

Se nos pide calcular $\boldsymbol{P(+c\ /\ -v, f_a, +t, -test)}$. De nuevo, usando las leyes y las propiedades de la probabilidad condicionada, concluimos que:

- $P(+c, -v, f_a, +t, -test) = P(-v) * P(+c/-v) * P(f_a/+c) * P(+t/+c) * $
  $* P(-test\ /\ +c) =$
  $= 0.5 * 0.7115 * 0.85 * 0.99 * 0.1 =$
  $= 0.0299363625$

- $P(-c, -v, f_a, +t, -test) = P(-v) * P(-c/-v) * P(f_a/-c) * P(+t/-c) * $
  $* P(-test\ /\ -c) =$
  $= 0.5 * (1 - 0.7115) * 0.005 * 0.01 * 0.99 =$
  $= 0.000007140375$

- $P(-v, f_a, +t, -test) = P(+c, -v, f_a, +t, -test) + P(-c, -v, f_a, +t, -test) =$
  $= 0.0299363625 + 0.000007140375 =$
  $= 0.029943502875$

- $\boldsymbol{P(+c\ /\ -v, f_a, +t, -test)} = P(+c, -v, f_a, +t, -test)\ /\ P(-v, f_a, +t, -test) =$
  $= 0.0299363625\ /\ 0.029943502875 =$
  $= \mathbf{0.9997615384202}$

Si comparamos los resultados de esta tarea con los obtenidos tras ejecutar la práctica 5.4 para las diferentes situaciones dadas, obtenemos lo siguiente:

- **Apartado 2:**
  - **Tarea:** $\boldsymbol{P(+c\ /\ z_a, -v, f_a, -t) = 0.9068089112017}$
  - **Práctica:** $\boldsymbol{P(+c\ /\ z_a, -v, f_a, -t) = 0.906808911201757}$
- **Apartado 3:**
  - **Tarea:** $\boldsymbol{P(-c\ /\ z_a, +v, f_a, +t) = 0.0058479532163}$
  - **Práctica:** $\boldsymbol{P(-c\ /\ z_a, +v, f_a, +t) = 0.005847953216374}$
- **Apartado 4:**
  - **Tarea:** $\boldsymbol{P(+c\ /\ -v, f_a, +t) = 0.9999975907284}$
  - **Práctica:** $\boldsymbol{P(+c\ /\ -v, f_a, +t) = 0.999975696011031}$
- **Apartado 5:**
  - **Tarea:** $\boldsymbol{P(+c\ /\ -v, f_a, +t, -test) = 0.9997615384202}$
  - **Práctica:** $\boldsymbol{P(+c\ /\ -v, f_a, +t, -test) = 0.999975696011031}$

Vemos que los resultados dados en cada apartado entre esta tarea y la práctica son casi idénticos. Esto es un buen indicador de que los cálculos se han realizado correctamente tanto en esta tarea como en la práctica.