

## 10 | Inferencias en Redes de Bayes usando factores

Verónica Esther Arriola Ríos

Esta práctica es complemento a la práctica [Factores](#).

### Objetivo

El alumno utilizará la clase factor para resolver consultas a una red de Bayes.

- Que el alumno identifique las ventajas de utilizar factores para calcular distribuciones de probabilidad sobre realizar los cálculos en la forma tradicional: con sumas y productos, considerando cada valor posible manualmente.
- Identificar la relación entre marginalización en probabilidad y marginalización de factores.
- Combinar las operaciones reducción y normalización de factores para calcular la reducción en distribuciones de probabilidad.
- Que el alumno pueda mapear de una ecuación con sumas y productos de distribuciones a las operaciones correspondientes con factores.

**Código Auxiliar 10.1:** Policías y ladrones

<https://github.com/computacion-ciencias/ia-policias-y-ladrones>

### Introducción

En esta práctica utilizarás un escenario de juguete donde un detective intenta resolver un robo. Listaremos a continuación la evidencia recabada por el detective, tu tarea será codificarla como un modelo gráfico probabilista y responder preguntas sobre el caso haciendo uso de inferencia.

## Antecedentes

Hace tres días el diamante *Colossus* fue robado del museo de *Ciudad Ladrillos*. El agente *Toño* está a cargo del caso y cuenta con la siguiente información:

Hay cinco personas sospechosas y se desea determinar quiénes participaron en el robo:

- Daniel
- Rita
- Joaquín
- Camelia
- Sam

de donde tendremos cinco variables aleatorias: DP, RP, JP, CP, SP, correspondiendo a la inicial del nombre y la palabra “participó”.

## Evidencia

1. Tanto Daniel como Joaquín suelen organizar robos de este tipo. De cada diez casos, Joaquín ha estado en siete y Daniel en cinco.
2. Rita ha participado anteriormente en robos con Daniel y Joaquín, aunque se lleva mejor con Daniel. Además, el empleo E de Rita estaba en riesgo, aunque el reporte no indica si lo perdió o no; la probabilidad de que lo hubiera perdido era del 65 %. Cuando Daniel y Joaquín participaron, ella estuvo con el equipo el 90 % de las ocasiones, si sólo estaba Daniel el 80 %, si sólo estaba Joaquín, el 45 %, y nunca actúa sin alguno de ellos. Si Rita está ocupada con su empleo esas probabilidades se reducen a la mitad.
3. Se sabe que Sam es amigo de Rita, y no le gusta el empleo de Rita. Por lo que, con tal de que lo deje, si ella conservó el empleo la probabilidad de que hubiera participado es del 80 % independientemente de que ella participara o no. Si Rita perdió el empleo y participó, la probabilidad de que Sam también estuviera involucrado es del 65 %, si ella no participó, es del 20 %.
4. El alarma del museo fue desactivada, esto se logra mediante el uso de alguno de tres dispositivos: A, B y C, de los cuales debieron haber elegido alguno ( $DiE$  = dispositivo elegido). Dependiendo de la pericia de su usuario, puede ser posible retirarlos antes de marcharse o no ( $DiA$  = dispositivo abandonado). Camelia trabaja en un laboratorio de electrónica donde se puede tener acceso a estos dispositivos y sabe muy bien cómo usarlos. Si ella participó existe una probabilidad del 20 % de que un dispositivo de tipo A haya sido abandonado en el museo, sino, esta probabilidad es del 35 %; ella dejará el dispositivo B con el 40 %, mientras que otro lo hará con un 85 %; sólo en un 10 % ella dejará un dispositivo C, mientras que otra persona lo dejará el 45 % de la veces.

5. Joaquín ha trabajado con Camelia en proyectos legales, existe una probabilidad del 15 % de que Camelia hubiera aceptado participar en el robo sólo si Joaquín lo hizo y un 1 % de que ella no hubiera participado y de hecho, ella hubiera ideado el robo. Si Joaquín no participó ella queda libre de toda sospecha pues no tendría ni conexión ni motivo para participar.
6. Independientemente de que Camelia haya estado ahí para usar el dispositivo, pudo haberse limitado a asesorar a los asaltantes (CA). Se sabe que Daniel la conoce y, de haber participado en el robo, hay un 50 % de probabilidades de que le haya pedido asesoría. Si no participó, no hubo quien solicitara la asesoría.
7. Entonces, el dispositivo elegido DiE para el robo dependió de si Camelia dió asesoría o participó el robo. Si Camelia participó en el robo definitivamente eligió el dispositivo C, pero si dió asesoría para que lo usara otra persona habrá sugerido el dispositivo B, si no participó ni asesorando ni en persona, un no especialista había elegido el dispositivo A el 25 % de las veces, el B el 40 % pues es más comercial y el C el 35 %.
8. La probabilidad de encontrar un dispositivo en el museo DiM depende del dispositivo que haya sido elegido y de que éste sea abandonado por quien lo usó. Si un dispositivo tipo A fue abandonado hay un 97 % de probabilidades de que la policía lo encuentre; 99 % para un dispositivo B y 90 % para un dispositivo C.
9. Para encontrar la ubicación U del diamante se concideran tres posibilidades:
  - Sam tiene una caja de seguridad en el banco, si él participó el diamante podría estar ahí.
  - Daniel es un mandón, pero algo imprudente, cuando ha participado en un robo frecuentemente los objetos han sido encontrados en su casa.
  - A Joaquín no le gusta correr riesgos y considera que Daniel es muy descuidado, pero también es muy metódico. En ocasiones anteriores se ha encontrado el botín enterrado en un punto equidistante a las casas de todos los participantes.

Si, de ellos tres, sólo Sam participó, el diamante estará en su caja de seguridad. Si Daniel participó, pero no Joaquín, hay un 80 % de probabilidades de que esté en casa de Daniel, un 15 % en la caja de Sam y un 5 % de que le hayan copiado la idea a Joaquín. Si Joaquín participó, pero no Daniel, hay un 65 % de probabilidad de que el diamante esté enterrado, un 30 % de que esté en la caja y un 5 % de que lo hayan sembrado en casa de Daniel. Si tanto Daniel como Joaquín estuvieron en el equipo, cualquiera de los tres lugares es igualmente probable... incluso si Sam no participó. Si ninguno de los tres participó la casa de Daniel sigue siendo un buen escondite con el 60 %, el punto central con 30 % y la caja de Sam 10 %.

## Desarrollo e implementación

Para resolver los ejercicios siguientes deberás utilizar la clase que programaste en la práctica Factores, por lo que será necesario que la entregues empaquetada en esta práctica también; si deseas hacer correcciones es un buen momento para incluirlas.

## Requisitos y resultados

Para las instrucciones siguientes asumiremos que el código está en `Python`. Si alguien decidió trabajar en otro lenguaje deberá consultar directamente con el ayudante del curso y adaptar las indicaciones a las convenciones del lenguaje elegido.

1. Para esta práctica deberás entregar dos archivos:

- a) `Factor.py`. Es el código de la práctica anterior y será necesario para ejecutar esta.
- b) `DetectiveFactor.py`. En este *script* colocarás las funciones para resolver los problemas indicados. Cada función que programes deberá:
  - 1) Imprimir al inicio cuál es la *consulta* (probabilidad) que está realizando.
  - 2) Los factores con resultados intermedios relevantes. Antes de imprimir el factor, imprime su significado. Por ejemplo:  $P(sp|E, RP)$ ,  $P(CP, jp)$  ó  $P(AD)P(U|E)$ .
  - 3) El resultado final, igualmente con algún formato que lo haga resaltar de las demás operaciones.

Al final, en la sección `if __name__ == '__main__':` mandar ejecutar todas las funciones mostrando tus resultados para cada ejercicio.

**Notación:** Para abreviar sumas/marginalizaciones sucesivas, se utilizará la siguiente abreviatura:

$$\sum_{A,B,C} = \sum_A \sum_B \sum_C f(A, B, C)$$

Recuerda usar paréntesis cuando las sumas no afecten a todos los términos a su derecha. Por ejemplo:

$$\left( \sum_A \sum_B f_1(A, B) \right) \left( \sum_C \sum_D f_2(C, D) \right)$$

Las tareas que debes realizar son las siguientes:

1. Dada la descripción anterior, genera **la gráfica de Bayes** con la información del caso así como **las tablas de distribución de probabilidad** correspondientes. Inclúyelas en un pdf, que entregarás con tu práctica.
2. Crea todas las variables correspondientes a las variables aleatorias como variables globales del *script*.
3. Crea todos los factores correspondientes a las distribuciones de probabilidad, también como variables globales.
4. Crea una función por cada ecuación, donde resuelvas directamente la traducción de las operaciones siguientes, en el orden en el que están especificadas. Para cada versión imprime el número de renglones que tienen tus factores después de cada multiplicación y marginalización. Observa que, al final, todas deben dar el mismo resultado

$$\begin{aligned}
 P(U) &= P(U|SP, JP, DP)P(SP|E, RP)P(RP|E, JP, DP)P(E)P(JP)P(DP) \quad (10.1) \\
 &= \sum_{SP, JP, DP, E, RP} P(U|SP, JP, DP)P(SP|E, RP)P(RP|E, JP, DP) \\
 &\quad P(E)P(JP)P(DP) \\
 &= \sum_E P(E) \sum_{RP, SP} P(SP|E, RP) \sum_{DP} P(DP) \sum_{JP} P(U|SP, JP, DP) \\
 &\quad P(RP|E, JP, DP)P(JP)
 \end{aligned}$$

5. Calcula la probabilidad de que:
  - Los cinco sospechosos hayan participado en el robo.
  - Ninguno de ellos haya sido.
6. Calcula la probabilidad de que Rita haya participado y un dispositivo tipo B haya sido abandonado.
7. Calcula la probabilidad de que Daniel haya participado, dado que se encontró el dispositivo B en el museo y el diamante estaba en la caja de seguridad de Sam. Para ello seguirás el método siguiente. Recuerda que, por definición de probabilidad condicional:

$$P(DP|D = A, U = Caja) = \frac{P(DP, D = A, U = Caja)}{P(D = A, U = Caja)}$$

- a) Calcula la ecuación correspondiente para obtener  $P(DP, D = A, U = Caja)$  usando la regla de la cadena para redes de Bayes. Factorízala según sea conveniente y escribe tu resultado en los comentarios de la función de python para este ejercicio.

- b) Utiliza tus factores para obtener el resultado correspondiente, pero antes de multiplicar y marginalizar, manda llamar la operación `reducir` sobre todos los factores que contengan a  $D$  y  $U$ , de modo que sólo te quedes con los renglones donde  $D = A$  y  $U = \text{Caja}$ .
- c) Normaliza tu resultado final, esta es la probabilidad que estabas buscando.