Inteligencia Artificial

Conocimientos y razonamientos bajo incertidumbre

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

Parte I

Cuantificando la Incertidumbre

En el que vemos cómo un agente puede manejar la incertidumbre usando grados de creencia

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

Actuando bajo incertidumbre

- Los agentes necesitan manejar la **incertidumbre** porque pueden desconocer **con seguridad** (observaciones parciales o falta de determinismo) su estado actual o cuál será después de llevar a cabo una secuencia de acciones.
- Los agentes estudiados hasta hora asumían que un agente se encontraba en un estado determinado
- Hay casos en que hay que elegir un plan de actuación en los que hay mucha incertidumbre y hay objetivos que cumplir

Ejemplo: Taxi automático

- Se trata de obtener un plan A_t donde t indica el tiempo con el que se ha de salir de casa
- Existe mucha incertidumbre sobre el estado del tráfico, el clima, y el propio estado del vehículo

"El plan A_{90} nos llevará al aeropuerto a tiempo, siempre y cuando el coche no se estropee o se quede sin gasolina, y yo no tenga un accidente, y no haya accidentes en el puente y el avión no sale temprano, y ningún meteorito golpea el coche, y. . . .

- ¿Por qué no A_{180}? La espera sería muy larga
- Lo correcto —la decisión racional— depende tanto de la importancia relativa de los diversos objetivos como de la probabilidad de que se logren y del grado en que se logren.

Intelligent Robotics Lab

Tomando decisiones racionales bajo incertidumbre

- En el plan del taxi A_{90}, podríamos tener un 97% de posibilidades de coger el vuelo
- Las posibilidad de A_{180} son mayores… y de A_{1440} aún más
- Para elegir, debemos tener preferencias sobre los resultados de los planes
- La **Teoría de la Utilidad** representa y razona sobre estas preferencias

La teoría de la utilidad dice que cada estado tiene un grado de utilidad para un agente y que el agente preferirá estados con mayor utilidad.

- Las preferencias, expresadas por las utilidades, se combinan con probabilidades en la teoría general de decisiones racionales llamada **teoría de la decisión**:

Teoría de la Decisión = Teoría de Probabilidad + Teoría de la Utilidad

Un agente es racional si y solo si elige la acción que produce la mayor utilidad esperada, promediada sobre todos los resultados posibles de la acción

Utilidad Máxima Esperada (MEU)

```
function DT-AGENT(percept) returns an action
  persistent: belief_state, probabilistic beliefs about the current state of the world
              action, the agent's action

  update belief_state based on action and percept
  calculate outcome probabilities for actions,
     given action descriptions and current belief_state
  select action with highest expected utility
     given probabilities of outcomes and utility information
  return action
```

Intelligent Robotics Lab

Ejemplo: diagnosticar el dolor de muelas

- Un dentista muy tajante diría

$$Toothache \Rightarrow Cavity$$

- Pero la verdad es que

$$Toothache \Rightarrow Cavity \vee GumProblem \vee Abscess$$

- Quizás podríamos decir

$$Cavity \Rightarrow Toothache$$

- Usar la lógica en este ejemplo no es apropiado debido a:
  - Es complejo tener una lista completa de antecedentes y consecuentes
  - Puede que aún haya razones desconocidas en el área
  - Puede no poder, o no ser necesario, hacer todos los tests al paciente

- La solución es asociar un **grado de certeza** a las sentencias relevantes, haciendo uso de la **teoría de probabilidad**

las opciones de $Toothache \Rightarrow Cavity$ podría ser un 80% (0.8 de probabilidad)

- ¿Cómo que es probable? En un mundo real, o se tiene una caries o no se tiene.
- Las declaraciones de probabilidad son sobre el conocimiento del estado, no sobre el mundo real

La probabilidad de que la paciente tenga caries, dado que tiene dolor de muelas, es de 0.8

La probabilidad de que la paciente tenga caries, dado que tiene dolor de muelas y antecedentes de enfermedad de las encías, es de 0.4 ← tenemos más conocimiento

Intelligent Robotics Lab

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

Notación de Probabilidad Básica

- **Espacio muestral** Ω: el conjunto de **todos los mundos posibles (los elementos del espacio muestral los llamamos ω)**
  - **Mutuamente exclusivo :** No puede haber dos mundos posibles iguales
  - **Exhaustivo :** Tienen que estar todos los mundos posibles.
  - Un **modelo de probabilidad** completamente especificado asocia una probabilidad numérica $P(\omega)$ con cada mundo posible.

$$0 \leq P(\omega) \leq 1 \text{ para cada } \omega \text{ y } \sum_{\omega \in \Omega} P(\omega) = 1 \quad (1)$$

- Ejemplo caso de tirar 2 dados (1,1) (1,2) .. Hay 36 posibilidades.

Por tanto Cada combinación de dados en Ω tiene probabilidad $\frac{1}{36}$

- Normalmente no se hacen preguntas ni aserciones sobre mundos posibles particulares, **se hacen sobre conjunto de mundos.** A esto se le llama **eventos**.

- Un **evento** es el conjunto, descrito como una **proposición en Lógica, de posibles mundos sobre los que se hace una aserción probabilística**, o una consulta. **Evento = Proposición (En lenguaje formal)**

$$\text{Para cualquier proposición } \phi,\ P(\phi) = \sum_{\omega \in \phi} P(\omega) \quad (2)$$

$$P(Total = 11) = P((5,6)) + P((6,5)) = \frac{1}{36} + \frac{1}{36} = \frac{1}{18}$$

$$P(Dobles) = P((1,1)) + P((2,2)) + P((3,3)) + P((4,4)) + P((5,5)) + P((6,6)) = 6 * \frac{1}{36} = \frac{1}{6} = 0.16666$$

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

- Las posibilidades como $P(Total = 11)$ o $P(Dobles)$ son **probabilidades incondicionales**. Se refieren a grados de creencia de proposiciones en ausencia de cualquier otra información.
- Pero en algunos casos tenemos **evidencias**, pudiendo construir **probabilidades condicionadas**

$$P(cavity) = 0.2$$

$$P(cavity \mid Toothache) = 0.6$$

Probabilidad condicionada

$$\boxed{P(a \mid b) = \frac{P(a \land b)}{P(b)}} \qquad (3)$$

$$P(doubles \mid Die_1 = 5) = \frac{P(doubles \land Die_1 = 5)}{P(Die_1 = 5)} = \frac{1}{6}$$

Regla del producto

$$\boxed{P(a \land b) = P(a \mid b)P(b)} \qquad (4)$$

Principio inclusivo-exclusivo

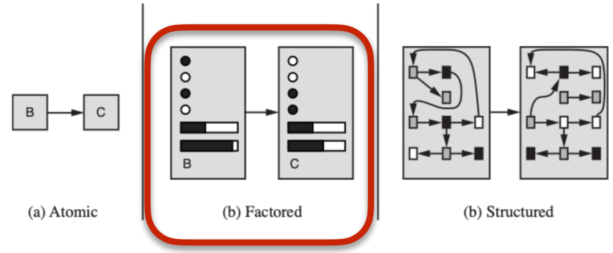
$$\boxed{P(a \lor b) = P(a) + P(b) - p(a \land b)} \qquad (5)$$

!

Ecuaciones 1 y 5 son llamadas axiomas de Kolgomorov

Intelligent Robotics Lab

- Las proposiciones que describen conjuntos de mundos posibles (eventos) están escritas en una notación que combina **elementos de lógica proposicional** y **notación de satisfacción de restricciones - representación factorizada**
- Las variables en probabilidad se llaman **variables aleatorias (Primera letra siempre en Mayúsculas)**, cada una con un **dominio** (valores que puede tomar). Cada una de ellas es una función Ω -> Dominio. (**Los valores siempre empiezan con minúsculas**)



Ejemplo de variables aleatorias:

Si no hay ambigüedad, se puede eliminar la variable $Weather = sunny$ se simplifica a $sunny$

- **Dominio de *Total*** en el caso de dos dados es $\{2,...,12\}$
- **Dominio de una variable aleatoria booleana** $\{true, false\}$

$A = True$ se abrevia como a
$A = False$ se abrevia como $\neg a$

- **Dominio de *Weather*** $\{sunny, rain, cloudy, snow\}$
- **Dominio de *Age*** $\{juvenile, teen, adult\}$

- Podemos combinar las proposiciones elementales usando lógica proposicional:

"La probabilidad de que el paciente tenga caries dado que es una adolescente sin dolor de dientes es de 0.1"

$$P(cavity \mid \neg toothache \wedge teen) = 0.1$$

- Podemos tener una **distribución de probabilidad**, para asociar una probabilidad con cada posible valor. En este caso se llama una **distribución categórica.**

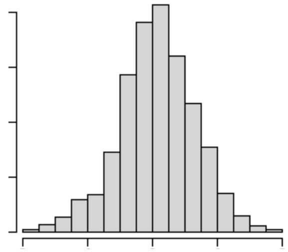
$$P(Weather = sunny) = 0.6$$
$$P(Weather = rain) = 0.1$$
$$P(Weather = cloudy) = 0.29$$
$$P(Weather = snow) = 0.01 \,,$$

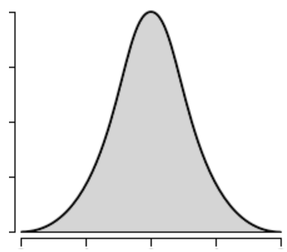
$$\mathbf{P}(Weather) = \langle 0.6, 0.1, 0.29, 0.01 \rangle$$

Intelligent Robotics Lab

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

- Para variables continuas usamos **funciones de densidad de probabilidad (PDFs).** En este ejemplo refleja la creencia de que la temperatura al mediodía se distribuye de forma uniforme entre 18 y 26 grados





$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x)$$

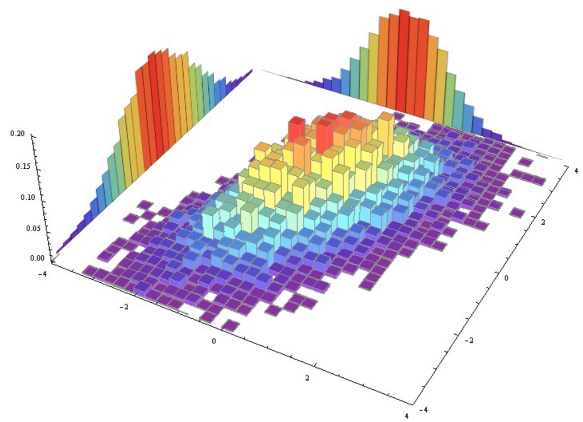
$$\boxed{P(x) = \lim_{dx \to 0} P(x \le X \le x + dx)/dx}$$

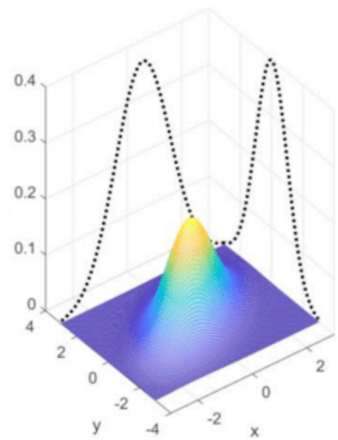
$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x) = \begin{cases} \frac{1}{8C} \text{ if } 18C \le x \le 26C \\ 0 \text{ otherwise} \end{cases}$$

- Para distribuciones discretas de más de una variable usamos **distribuciones de probabilidad conjuntas**

$$\mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather \mid Cavity)\mathbf{P}(Cavity)$$

$$\begin{aligned}
&P(W = sunny \land C = true) = P(W = sunny|C = true)\, P(C = true)\\
&P(W = rain \land C = true) = P(W = rain|C = true)\, P(C = true)\\
&P(W = cloudy \land C = true) = P(W = cloudy|C = true)\, P(C = true)\\
&P(W = snow \land C = true) = P(W = snow|C = true)\, P(C = true)\\
&P(W = sunny \land C = false) = P(W = sunny|C = false)\, P(C = false)\\
&P(W = rain \land C = false) = P(W = rain|C = false)\, P(C = false)\\
&P(W = cloudy \land C = false) = P(W = cloudy|C = false)\, P(C = false)\\
&P(W = snow \land C = false) = P(W = snow|C = false)\, P(C = false)\,.
\end{aligned}$$





- Una **distribución de probabilidad conjunta completa** es la distribución conjunta de todas las variables aleatorias, y determina completamente un modelo de probabilidad. (2x2x4 tabla con 16 entradas)

Intelligent Robotics Lab

Inferencia usando distribuciones conjuntas completas

- Vamos a hacer **inferencia probabilística** usando las **distribuciones conjuntas completas** como **base de conocimiento**.

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

$$\sum_{\omega \in \Omega} P(\omega) = 1$$

Catch: la desagradable sonda de acero del dentista se atora en mi diente

Para una proposición ϕ, $P(\phi) = \sum_{\omega \in \phi} P(\omega)$

$$P(cavity \vee toothache) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$$

- **Probabilidad marginal : Tarea común de extraer la distribución sobre algún subconjunto de variables o una variable**
- **Probabilidad marginal de** $cavity$

$$P(cavity) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$$

$$\mathbf{P}(Cavity) = \sum_{z \in \{Catch, Toothache\}} P(Cavity, z)$$

Intelligent Robotics Lab

Inferencia usando distribuciones conjuntas completas

Regla del producto (4)

Marginalización $\mathbf{P}(Y) = \sum_{z \in Z} P(Y, z)$ → Regla de Condicionamiento $\mathbf{P}(Y) = \sum_{z} P(Y \mid z) P(z)$

- En multitud de ocasiones estamos interesados en calcular las probabilidades condicionales de algunas variables, dadas evidencias sobre otras.

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Probabilidad de Caries dada la evidencia de que existe dolor de muelas

$$P(cavity \mid toothache) = \frac{P(cavity \land toothache)}{P(toothache)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

$$P(\neg cavity \mid toothache) = \frac{P(\neg cavity \land toothache)}{P(toothache)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

vector

todos los posibles casos.

lo estamos poniendo a capón por lo tanto podemos normalizar.

$\alpha = \frac{1}{P(toothache)}$ ← te olvidas del dolor de muelas.

$$\begin{aligned}\mathbf{P}(Cavity \mid toothache) &= \alpha\, \mathbf{P}(Cavity, toothache) \\ &= \alpha\, [\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)] \\ &= \alpha\, [\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle] = \alpha\, \langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle\,.\end{aligned}$$

Normalización α

- **Procedimiento de inferencia general**
  - **La consulta** implica una variable X (por ejemplo, *Cavity*)
  - E es la lista de **evidencias** (*Toothache*, por ejemplo), siendo e un valor observado
  - Y **es el resto de variables no observadas** (*Catch*, por ejemplo)

$$\boxed{\mathbf{P}(X \mid e) = \alpha \sum_{y} \mathbf{P}(X, e, y)} \qquad (9)$$

Dada una distribución conjunta completa para trabajar, la ecuación 9 puede responder a preguntas probabilísticas sobre variables discretas. Pero no escala bien. Por esta razón **este método no se utiliza** generalmente para construir sistemas de razonamiento. Es la base teórica sobre la que se construyen algoritmos más eficientes. !

Intelligent Robotics Lab

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

Independencia

- Imaginemos que añadimos a nuestra distribución conjunta completa

$$\mathbf{P}(Toothache, Catch, Cavity, Weather)$$

$$\begin{aligned} &P(toothache, catch, cavity, cloudy) \\ &= P(cloudy \mid toothache, catch, cavity) P(toothache, catch, cavity) \end{aligned}$$

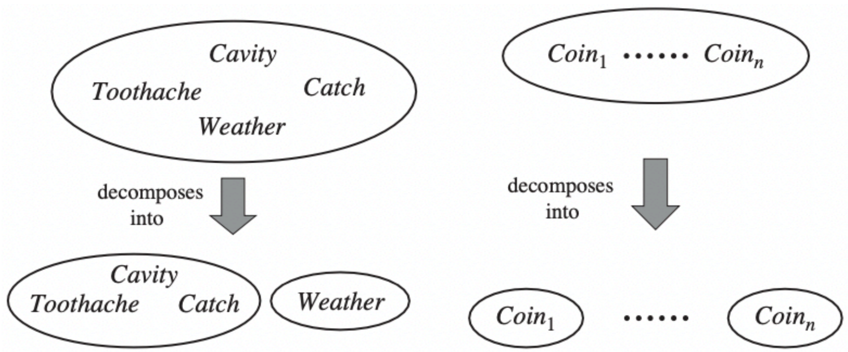
Al no depender $cloudy$ de las otras variables $P(cloudy \mid toothache, catch, cavity) = P(cloudy)$

$$P(toothache, catch, cavity, cloudy) = P(cloudy) P(toothache, catch, cavity)$$

$$\boxed{P(a \mid b) = P(a) \quad \text{or} \quad P(b \mid a) = P(b) \quad \text{or} \quad P(a \wedge b) = P(a)P(b)}$$

$$\boxed{\mathbf{P}(X \mid Y) = \mathbf{P}(X) \quad \text{or} \quad \mathbf{P}(Y \mid X) = \mathbf{P}(Y) \quad \text{or} \quad \mathbf{P}(X, Y) = \mathbf{P}(X)\mathbf{P}(Y)}$$

Independencia



Ejercicio 1

Given the full joint distribution shown in Figure 13.3, calculate the following:

1. $\mathbf{P}(toothache)$.
2. $\mathbf{P}(Catch)$.
3. $\mathbf{P}(Cavitycatch)$.
4. $\mathbf{P}(Cavitytoothache \vee catch)$.

Intelligent Robotics Lab

La Regla de Bayes y su uso

"Las afirmaciones extraordinarias requieren evidencias extraordinarias"
-Carl Sagan-

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

La Regla de Bayes y su uso

- $\boxed{\text{P(hipótesis|datos) = P(hipótesis) x P(datos | hipótesis) / P (datos)}}$

P(hipótesis | datos) : probabilidad a posteriori. Nuestra creencia actualiza la hipótesis una vez examinadas las evidencias. (Ejemplo: nuestra confianza en el diagnóstico de una enfermedad una vez vistos los resultados)

P(hipótesis) : Probabilidad previa o a priori. Nuestra creencia en la hipótesis antes de analizar los datos. (En el caso de una enfermedad esta podría ser su prevalencia en la población, **la tasa base**)

P(datos | hipótesis) : **Se denomina verosimilitud (likelihood)**, no se refiere a probabilidad, sino que se refiere a cuán probables serían los datos si la hipótesis fuera verdadera. Si alguien tiene la enfermedad, ¿Cómo de probable es que muestre un síntoma dado o que dé positivo en una prueba?

P(datos) : probabilidad de los datos en general, tanto si la hipótesis es verdadera como falsa. A veces se la designa como **probabilidad marginal**. En el caso del diagnóstico médico se refiere a la proporción de todos los pacientes que tienen un síntoma u obtienen un resultado positivo, tanto sanos como enfermos.

ejemplo un hipocondriaco

En lenguaje común : La probabilidad a posteriori = probabilidad a priori x verosimilitud de los datos / habitualidad de los datos

“Ahora que he visto las evidencias, ¿cuánto debería de creer en la idea?”

!
+ Si está bien respaldada de entrada (primer término -tasa base- del numerador alto)
+ Si es probable que ocurran las evidencias cuando la idea es verdadera (verosimilitud)
- Si las evidencias son habituales (denominador), si tiene una alta probabilidad marginal.

Pinker, S. and Lazcano, P. H. “Racionalidad: Qué es, por qué escasea y cómo promoverla” 2021, Ediciones Paidós pp 186-188

Intelligent Robotics Lab

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

La Regla de Bayes y su uso

Ejemplo de cáncer de mama:

Supongamos que la prevalencia del cáncer de mama en la población de mujeres es del 1%. Supongamos que la sensibilidad de una prueba (positivos verdaderos) es del 90%. Supongamos que los falsos positivos son el 9%. **Una mujer da positivo en el test. ¿Cuál es la probabilidad de que tenga la enfermedad?**

Prevalencia en la población 1% (tasa base) -> P(hipótesis) = 0.01
La sensibilidad de la prueba (verosimilitud) -> es la posibilidad de obtener resultados positivos dado que el paciente tiene la enfermedad. 90% . P(datos|hipótesis) = 0,9
Probabilidad marginal del resultado positivo en una prueba general es la suma de las probabilidades de un acierto para los pacientes enfermos (el 90% del 1% o 0,009) y de una falsa alarma para los sanos (9% del 99% o 0,0891 .. 0,1)

Bayes nos da 0,01 *0,9/0,1 = 0,09 **9% de probabilidades de tener la enfermedad.**

Pinker, S. and Lazcano, P. H. “Racionalidad: Qué es, por qué escasea y cómo promoverla” 2021, Ediciones Paidós. Pp 184

Elementary Bayesian inference

Example from Daniel Kahneman's *Thinking Fast and Slow*

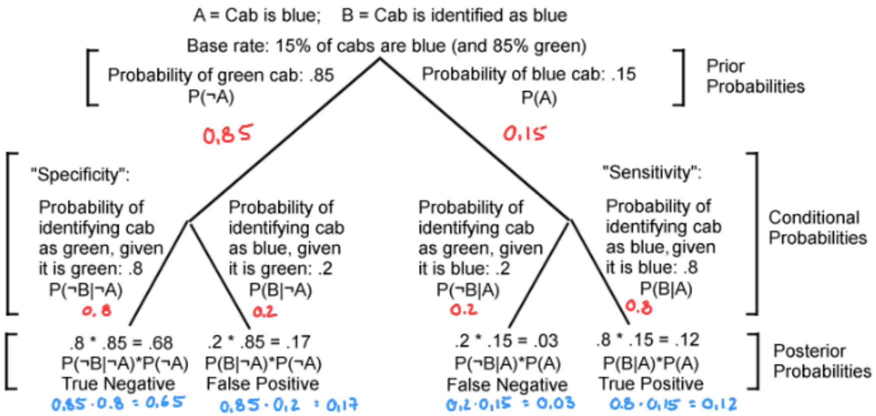
For my grandkids:

In his book Thinking Fast and Slow, Daniel Kahneman gives an example of elementary Bayesian inference, posing this question:

> "A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data: 85% of the cabs in the city are Green and 15% are Blue. A witness identified the cab as Blue. The court tested the reliability of the witness under the circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time. What is the probability that the cab involved in the accident was Blue rather than Green?"

Kahneman goes on to observe that "The two sources of information can be combined by Bayes's rule. The correct answer is 41%. However, you can probably guess what people do when faced with this problem: they ignore the base rate and go with the witness. The most common answer is 80%."

So why is the correct answer 41%?



So what is the probability that the cab involved in the accident really was blue, given that it was identified as blue? This is the posterior probability P(A|B): the true positive probability P(B|A)*P(A) divided by the probability of the cab being identified as blue (P(B)), which, in turn, is the sum of the true positive and false positive probabilities, that is, P(B|A)*P(A) + P(B|¬A)*P(¬A).

$$\boxed{P(A|B) =} \frac{P(B|A)*P(A)}{P(B)} \quad \left(\text{The usual statement of Bayes' Theorem}\right)$$

$$= \frac{P(B|A)*P(A)}{P(B|A)*P(A) + P(B|\neg A)*P(\neg A)}$$

pasos positivos como negativos.

$$= \frac{.12}{.12 + .17}$$

$$= 0.413793103$$

$$= \boxed{\text{About } 41\%}$$

Which is the answer given by Kahneman, et quod erat demonstrandum

https://www.anesi.com/bayes.htm?p_a=0.15&p_b_a=0.8&p_not_b_not_a=0.8

Intelligent Robotics Lab

La Regla de Bayes y su uso

$$\boxed{P(b \mid a) = \frac{P(a \mid b)P(b)}{P(a)}}$$

- En caso de variables multivalor, y con el conocimiento de alguna evidencia e

$$\mathbf{P}(Y \mid X) = \frac{\mathbf{P}(X \mid Y)\mathbf{P}(Y)}{\mathbf{P}(X)},$$

- Puede que la regla de Bayes no parezca muy útil, pero lo es $\mathbf{P}(Y \mid X, \mathbf{e}) = \frac{\mathbf{P}(X \mid Y, \mathbf{e})\mathbf{P}(Y \mid \mathbf{e})}{\mathbf{P}(X \mid \mathbf{e})}.$

$$P(cause \mid effect) = \frac{P(effect \mid cause)P(cause)}{P(effect)}$$

- P(effect|cause) -> Cuantifica la relación en la dirección causal. El doctor conoce:

$$P(symptoms \mid disease) \leftarrow$$ Teniendo la enfermedad, conocer los síntomas.

- P(cause|effect) -> Cuantifica la relación en la dirección del diagnóstico. El doctor quiere derivar:

$$P(disease \mid symptoms) \leftarrow$$ Teniendo los síntomas, conocemos la enfermedad.

Intelligent Robotics Lab

La Regla de Bayes y su uso

- Veamos otro ejemplo:
  - Un médico sabe que la enfermedad meningitis hace que el paciente tenga rigidez en el cuello, digamos, el 70% de las veces
  - El médico también conoce algunos hechos incondicionales:
    - (Tasa base) la probabilidad previa de que un paciente tenga meningitis es 1 / 50.000 y
    - (Probabilidad marginal) la probabilidad previa de que cualquier paciente tenga rigidez de cuello es del 1%.
  - Siendo ***s*** la proposición de que el paciente tiene rigidez en el cuello y ***m*** la proposición de que el paciente tiene meningitis,

Examen :

$$\begin{aligned} P(s \mid m) &= 0.7 \\ P(m) &= 1/50000 \leftarrow \text{tasa base (prevalencia en la población sin tener en cuenta nada más)} \\ P(s) &= 0.01 \leftarrow \text{probabilidad marginal (la tenemos todos) dolor de cuello.} \\ P(m \mid s) &= \frac{P(s \mid m)P(m)}{P(s)} = \frac{0.7 \times 1/50000}{0.01} = 0.0014 \end{aligned}$$

$$\mathbf{P}(M \mid s) = \alpha \, \langle P(s \mid m)P(m), P(s \mid \neg m)P(\neg m) \rangle$$

No necesitamos $p(s)$

Intelligent Robotics Lab

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- ¿Qué pasa cuando tenemos más de una evidencia? Anteriormente, hemos descrito como calcular

$$\mathbf{P}(Cavity \mid toothache \land catch) = \alpha \, \langle 0.108, 0.016 \rangle \approx \langle 0.871, 0.129 \rangle$$

Cavity

¬Cavity

- Puede que no escale bien para un gran número de variables. Usemos Bayes:

lo sustituimos como variables independientes

$$\mathbf{P}(Cavity \mid toothache \land catch)$$
$$= \alpha \, \mathbf{P}(toothache \land catch \mid Cavity) \, \mathbf{P}(Cavity)$$

normalizamos, eliminamos el denominador.

calculamos lo + y −

- Puede que no escale bien tampoco si hay muchas variables de evidencias. Veamos cómo puede ayudarnos el concepto de independencia

$$\mathbf{P}(toothache \land catch \mid Cavity) = \mathbf{P}(toothache \mid Cavity)\mathbf{P}(catch \mid Cavity)$$

Independencia condicional

$$\boxed{\mathbf{P}(X, Y \mid Z) = \mathbf{P}(X \mid Z)\mathbf{P}(Y \mid Z)}$$

$$\mathbf{P}(Cavity \mid toothache \land catch)$$
$$= \alpha \, \mathbf{P}(toothache \mid Cavity) \, \mathbf{P}(catch \mid Cavity) \, \mathbf{P}(Cavity)$$

$$\mathbf{P}(Cause, Effect_1, \ldots, Effect_n) = \mathbf{P}(Cause) \prod_i \mathbf{P}(Effect_i \mid Cause)$$

ingenuo

Bayes ingenuo (naive Bayes)

problema, la independencia.

Estamos sacrificando la precisión por la rapidez.

Intelligent Robotics Lab

Ejercicio 2

Los extraterrestres pueden ser amistosos o no; El 75% son amigables. Los extraterrestres amistosos llegan durante el día el 90% del tiempo, mientras que los hostiles siempre llegan por la noche. Si un extraterrestre llega de noche, ¿qué probabilidad hay de que sea amistoso?

A = amistosos
H = ¬A = hostiles

D = día
N = ¬D = noche

Fotograma de la película de Tim Burton 'Mars Attacks'.

$P(A) = 0{,}75$

$P(D|A) = 0{,}9$
$P(N|A) = 0{,}1$
$P(D|H) = 0$
$P(N|H) = 1$

$$P(A|N) = \frac{P(N|A) \cdot P(A)}{P(N)} = 0{,}23$$

Resolver.

Intelligent Robotics Lab

Alejandro Bascuñana Muñoz **alejandro.bascunana@urjc.es**

Bibliografía

Capítulo 12. Russell S., Norving P. “Artificial intelligence, A Modern Approach”. 2022, Pearson.

Capítulo 5. Pinker, S. and Lazcano, P. H. “Racionalidad: Qué es, por qué escasea y cómo promoverla” 2021, Ediciones Paidós.

Alejandro Bascuñana Muñoz alejandro.bascunana@urjc.es