

Arboles de decisión

Teoría de la decisión

- ✍ Hemos estudiado las redes bayesianas como instrumento para el diagnóstico en presencia de incertidumbre.
- ✍ La idea ahora es aprovechar ese diagnóstico para decidir sobre el tratamiento a realizar
- ✍ En realidad estudiaremos la toma de decisiones en presencia de incertidumbre.

Valor esperado

- ✍ **Nuestro punto de partida es el concepto de valor esperado, aplicable a *variables aleatorias que toman valores numéricos*.**
- ✍ **Intuitivamente el valor esperado es el promedio de los valores que se obtendrían al realizar el experimento un número muy grande de veces.**
- ✍ **Es decir, es el valor al que tiende el promedio.**

$$E[X] = \sum_x x \cdot P(x)$$

- ✍ **Cuando todos los valores son igualmente probables el valor esperado es simplemente el promedio de los valores posibles, pues si hay n valores $P(x) = 1/n$ y por lo tanto**

$$E[X] = \sum_x x \cdot \frac{1}{n}$$

Utilidad esperada

- ✍ Si cada valor representa un beneficio o una pérdida podemos considerarlo como una **utilidad**; cuanto más positiva es la utilidad, mayor es el beneficio, y cuanto más negativa es mayor es el perjuicio.
- ✍ Por lo tanto da una medida para tomar decisiones: elegiremos la opción que tenga mayor utilidad (*principio de máxima utilidad*)
- ✍ En general, si la utilidad es función de una variable aleatoria X , la utilidad esperada viene dada por:

$$UE = \sum_x U(x) P(x)$$

- ✍ De este modo la teoría de la probabilidad nos conduce a la **teoría de la decisión**.

Utilidad esperada

- ✍ Existen dos enfoques en cuanto a la consideración de la utilidad:
 - ✍ Algunos afirman que el mejor juego era aquél que tuviera el mayor valor económico (*principio de la esperanza matemática*)
 - ✍ Bernouilli (1738) introdujo el término de utilidad como la satisfacción que experimenta un sujeto al recibir unan ganancia o una pérdida, distinguiendo así entre **valor (objetivo) y utilidad (subjetiva)**.
 - ✍ Esto hace que distintas personas tengan actitudes diferentes ante el riesgo: unas personas están más dispuestas a arriesgarse en cuestiones dominadas por el azar, mientras que otras hacen todo lo posible por evitar el riesgo.

Diagramas de influencia

- ✍ **Análogamente a lo que ocurre con las redes bayesianas, un diagrama de influencia viene dado por un grafo y unas tablas.**
- ✍ **En el grafo pueden aparecer tres tipos de nodos:**
 - ✍ **Variables aleatorias**, que se representan por círculos
 - ✍ **Decisiones**, que se representan por cuadrados o rectángulos
 - ✍ **Utilidad**, que se representan por rombos.
- ✍ **Cada nodo aleatorio lleva asociado una tabla de probabilidad condicional dados sus padres (que pueden ser tanto aleatorios como de decisión)**

Diagramas de influencia

- ✍ El significado de cada enlace viene dado por la naturaleza de los nodos que une:

Destino/Origen	Variable aleatoria X	Decisión D1
Variable aleatoria Y	X influye causalmente de forma directa sobre Y	La decisión D1 influye directamente sobre Y
Decisión D2	Al decidir D2 se conoce ya el valor que ha tomado X	La decisión D1 se toma antes que D2
Utilidad U	El valor que toma X influye directamente en la utilidad	La decisión D1 influye directamente en la utilidad

Arboles de decisión

✍ Se compone de:

- ✍ Un nodo raíz, que puede ser aleatorio o de decisión
- ✍ Cada nodo (excepto el nodo de utilidad) tiene varios hijos, uno por cada valor de la variable asociada al nodo
- ✍ Gráficamente el nodo raíz se representa a la izquierda, y los hijos de cada nodo a la derecha de su padre, siguiendo un orden temporal en las acciones; por tanto los nodos de utilidad aparecen en el extremo de la derecha.
- ✍ Los nodos aleatorios se representan por círculos, y los de decisión por cuadrados.
- ✍ Los enlaces que parten de un nodo aleatorio X llevan asociada una probabilidad condicional $P(x|izq(x))$, donde $izq(x)$ representa los valores que toman los nodos situados a la izquierda de X en el árbol.

De un diagrama de influencia a un árbol de decisión

- ✍ **Para representar un árbol de decisión se deben seguir las siguientes reglas:**
 - ✍ **Regla 1:** A la derecha del todo debe aparecer el valor de la utilidad para cada rama.
 - ✍ **Regla 2:** Si la Decisión D1 se toma antes que D2, el nodo D1 debe aparecer a la izquierda del nodo D2.
 - ✍ **Regla 3:** Las variables cuyo valor se conoce antes de tomar la decisión D, han de aparecer a la izquierda del nodo D en el desarrollo del árbol; las que no se conocen al decidir D, aparecerán a la derecha.
 - ✍ **Regla 4:** De las dos reglas anteriores se deduce que, si el valor de la variable X se conoce después de la Decisión D1 y antes de la D2, el nodo X debe aparecer entre D1 y D2.

Evaluación de un árbol de decisión

- ✍ La evaluación de un árbol de decisión se realiza siempre de derecha a izquierda.
- ✍ La utilidad asociada a cada rama y a cada nodo se calcula teniendo en cuenta que:
 - ✍ La utilidad correspondiente a un nodo aleatorio X en el árbol de decisión es el promedio de todas las ramas que parten de X, ponderado por la probabilidad (recordemos que hay una rama por cada valor x de X):

$$U_x(\text{izq}, x) = \sum_x U(x | \text{izq}, x) \cdot P(x | \text{izq}, x)$$

- ✍ La utilidad correspondiente a un nodo de decisión D es el máximo de las utilidades de sus ramas
 - ✍ La decisión óptima para este nodo es el valor D correspondiente a la rama de mayor utilidad.

$$U_D(\text{izq}, d) = \max_d U(d | \text{izq}, d)$$