

# Aprendizaje automatizado

## MODELOS GRÁFICOS PROBABILISTAS

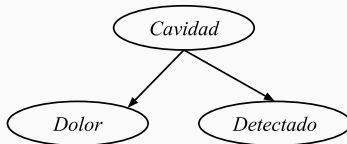
---

Gibran Fuentes-Pineda

Marzo-Abril 2023

# Distribución conjunta completa

- Contiene toda la información necesaria para obtener cualquier probabilidad.
- Por ejemplo



	<i>Do</i> = Sí		<i>Do</i> = No	
	<i>De</i> = Sí	<i>De</i> = No	<i>De</i> = Sí	<i>De</i> = No
<i>Ca</i> = Sí	0.108	0.012	0.072	0.008
<i>Ca</i> = No	0.016	0.064	0.144	0.576

Ejemplo de AIMA (Russel y Norvig 2009)

# Cálculo de probabilidades marginales

- Para obtener una probabilidad marginal a partir de las probabilidades conjuntas aplicamos la regla de la suma.
- Por ejemplo

$$P(Ca = \text{Sí}) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$$

	<i>Do = Sí</i>		<i>Do = No</i>	
	<i>De = Sí</i>	<i>De = No</i>	<i>De = Sí</i>	<i>De = No</i>
<i>Ca = Sí</i>	0.108	0.012	0.072	0.008
<i>Ca = No</i>	0.016	0.064	0.144	0.576

Ejemplo de AIMA (Russel y Norvig 2009)

# Cálculo de probabilidades condicionales

- Para obtener una probabilidad condicional a partir de la distribución conjunta aplicamos la regla del producto.
- Por ejemplo

$$\begin{aligned}P(Ca = \text{Sí} | Do = \text{Sí}) &= \frac{P(Ca = \text{Sí}, Do = \text{Sí})}{P(Do = \text{Sí})} \\&= \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.2\end{aligned}$$

	<i>Do</i> = Sí		<i>Do</i> = No	
	<i>De</i> = Sí	<i>De</i> = No	<i>De</i> = Sí	<i>De</i> = No
<i>Ca</i> = Sí	0.108	0.012	0.072	0.008
<i>Ca</i> = No	0.016	0.064	0.144	0.576

Ejemplo de AIMA (Russel y Norvig 2009)

## Número de parámetros para distribución conjunta

- ¿Cuántas probabilidades tenemos que calcular si agregamos la variable discreta  $T_e$  con 10 posibles valores?

## Número de parámetros para distribución conjunta

- ¿Cuántas probabilidades tenemos que calcular si agregamos la variable discreta  $T_e$  con 10 posibles valores?
- Crecimiento exponencial: por ej. para  $n$  variables discretas de  $K$  valores sería  $K^n$

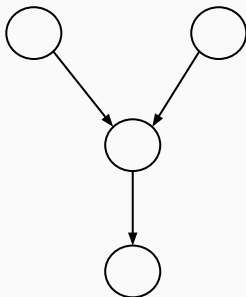
# Número de parámetros para distribución conjunta

- ¿Cuántas probabilidades tenemos que calcular si agregamos la variable discreta  $T_e$  con 10 posibles valores?
- Crecimiento exponencial: por ej. para  $n$  variables discretas de  $K$  valores sería  $K^n$
- Los modelos gráficos reducen la complejidad factorizando la distribución conjunta en distribuciones condicionales y aprovechando las relaciones de independencia

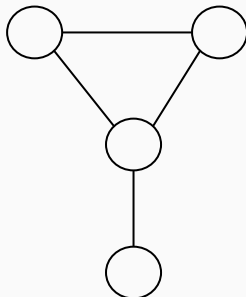
# Modelos gráficos probabilistas

- Define una familia de distribuciones conjuntas de probabilidad sobre un conjunto de variables aleatorias

Grafo dirigido



Grafo no dirigido





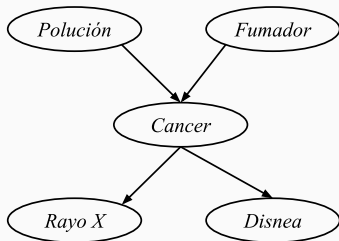
- **Nodos** representan variables aleatorias
- **Aristas** representan dependencias entre variables aleatorias
- La red representa relaciones causa-efecto
- Aprovecha independencias para especificar de forma compacta la distribución conjunta completa
  - Las aristas corresponden a dependencias directas entre variables
  - La ausencia de aristas captura las independencias absolutas y condicionales

- **Representación:** especificación de variables aleatorias y sus dependencias e independencias
- **Inferencia:** consulta de probabilidades en la red dado el modelo y ciertas evidencias
- **Aprendizaje:** obtener la topología y/o los parámetros de las distribuciones de la red a partir de ejemplos

- Estructura del grafo: variables aleatorias y dependencias e independencias
- Distribución conjunta se expresa por la distribución condicional de cada nodo dados sus padres (factorización)

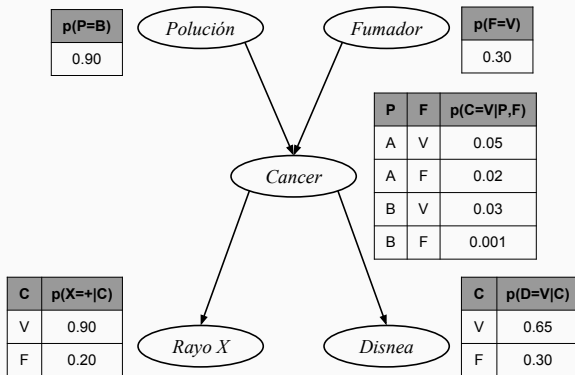
$$p(x_1, \dots, x_d) = \prod_{i=1}^d p(x_i | \text{Padres}(x_i))$$

## Ejemplo: cáncer



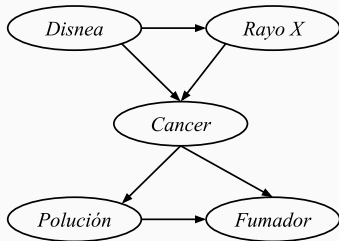
Ejemplo de BAI (Korb y Nicholson 2010)

# Ejemplo: cáncer



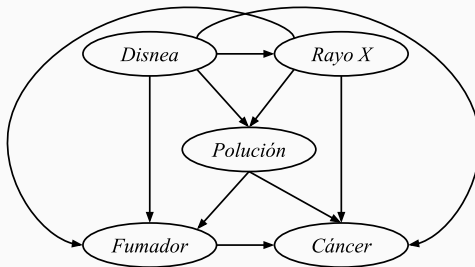
Ejemplo de BAI (Korb y Nicholson 2010)

## Ejemplo: cáncer visto de otra manera



Ejemplo de BAI (Korb y Nicholson 2010)

## Ejemplo: cáncer visto de otra manera



Ejemplo de BAI (Korb y Nicholson 2010)

- **Mapa-I:** No hay dependencias directas en la distribución conjunta que no estén especificadas en el grafo explícitamente
- **Mapa-D:** Todas las aristas en el grafo corresponden a dependencias directas en la distribución conjunta
- **Mapa perfecto:** Modelo gráfico con mapa-I y mapa-D



## Recordando la independencia condicional

- La variable aleatoria  $x$  es independiente de  $y$  dado  $z$  si

$$P(x, y|z) = P(x|z)P(y|z)$$

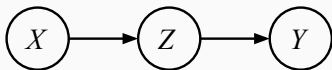
$$P(x|y, z) = P(x|z)$$

- La independencia condicional se denota con el símbolo  $\perp\!\!\!\perp$

$$x \perp\!\!\!\perp y \mid z$$

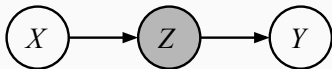
## Independencia condicional: cadenas causales

$$x \not\perp y \mid \emptyset$$



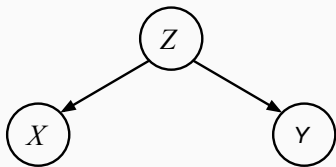
## Independencia condicional: cadenas causales

$$x \perp\!\!\!\perp y \mid z$$



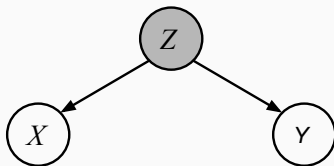
## Independencia condicional: causas comunes

$$X \not\perp Y \mid \emptyset$$



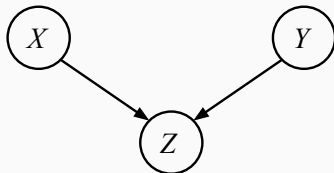
## Independencia condicional: causas comunes

$$x \perp\!\!\!\perp y \mid z$$



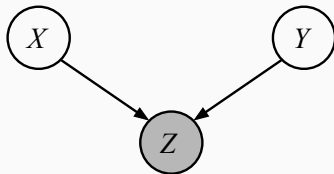
## Dependencia condicional: efectos comunes

$$x \perp\!\!\!\perp y \mid \emptyset$$



## Dependencia condicional: efectos comunes

$$x \not\perp\!\!\!\perp y \mid z$$



- Secuencia de nodos entre un miembro de  $X$  y un miembro de  $Y$  tal que cada par de nodos adyacente está conectado por una arista sin importar la dirección



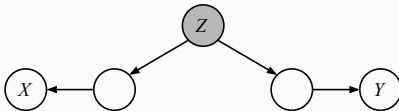
# Caminos bloqueados

- Dado un conjunto de nodos  $Z$ , se dice que el camino está bloqueado si hay un nodo  $z$  para el cual se cumple al menos que
  - $z$  está en  $Z$  y tiene una arista sobre el camino entra y la otra sale (cadena)



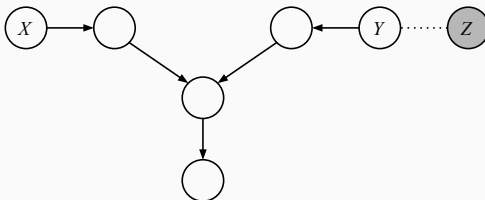
# Caminos bloqueados

- Dado un conjunto de nodos  $Z$ , se dice que el camino está bloqueado si hay un nodo  $z$  para el cual se cumple al menos que
  - $z$  está en  $Z$  y tiene una arista sobre el camino entra y la otra sale (cadena)
  - $z$  está en  $Z$  y ambas aristas salen (causa común)



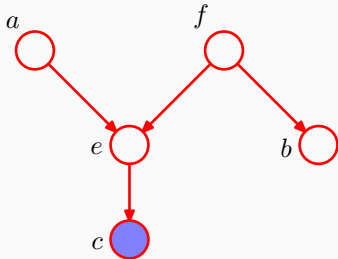
# Caminos bloqueados

- Dado un conjunto de nodos  $Z$ , se dice que el camino está bloqueado si hay un nodo  $z$  para el cual se cumple al menos que
  - $z$  está en  $Z$  y tiene una arista sobre el camino entra y la otra sale (cadena)
  - $z$  está en  $Z$  y ambas aristas salen (causa común)
  - $z$  ni sus descendientes está en  $Z$  y ambos caminos entran a  $z$  (efecto común)



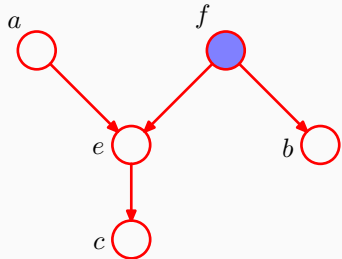
# Separación D: ejemplo

$$a \not\perp b \mid c$$



Tomada de PRML (Bishop 2009)

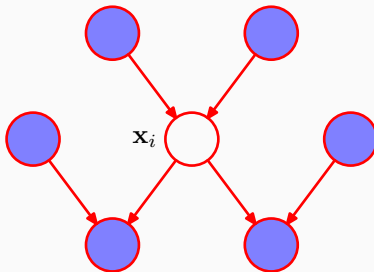
$$a \perp b \mid f$$



Tomada de PRML (Bishop 2009)

- Conjunto de nodos  $Y$  que hacen un nodo  $x$  independiente de cualquier otro nodo  $z$  en el grafo: padres, hijos y otros padres de hijos

$$P(x|Y, z) = P(x|Y)$$

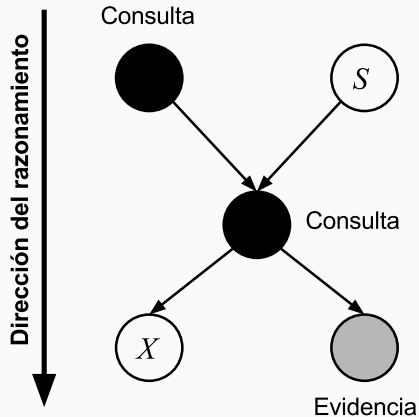


Tomada de PRML (Bishop 2009)

- Exacta
  - Intratable para distribuciones y topologías generales
  - Eficiente para algunas topologías (por ej. árboles, poliárboles, etc.)
  - Eliminación de variables, propagación de creencias, etc.
- Aproximada
  - Muestreo directo
  - Muestreo por rechazo
  - Pesado de verosimilitud
  - Montecarlo por cadenas de markov

# Tipos de inferencia en redes bayesianas

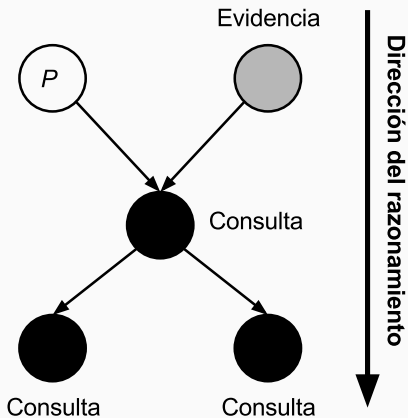
- Diagnóstico



Basada en figura de BAI (Korb y Nicholson 2010)

# Tipos de inferencia en redes bayesianas

- Predictiva

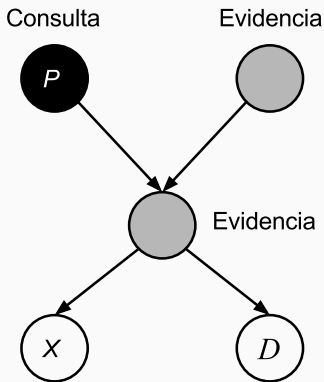


Basada en figura de BAI (Korb y Nicholson 2010)



# Tipos de inferencia en redes bayesianas

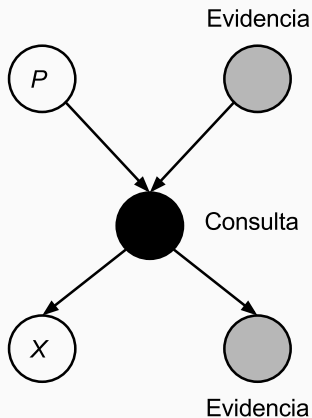
- Intercausal (justificación)



Basada en figura de BAI (Korb y Nicholson 2010)

# Tipos de inferencia en redes bayesianas

- Combinada



Basada en figura de BAI (Korb y Nicholson 2010)

- **Específica:** la variable toma un valor particular (por ej. el paciente es fumador)

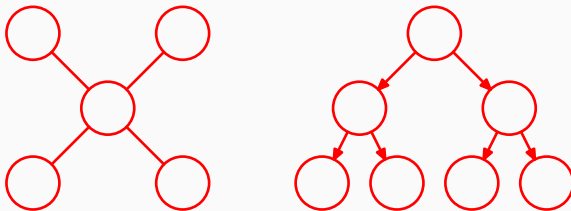
- **Específica:** la variable toma un valor particular (por ej. el paciente es fumador)
- **Negativa:** la variable puede tomar un subconjunto de valores, descartando los demás (por ej. la polución no es alta o la polución es baja o media)

# Tipos de evidencia

- **Específica:** la variable toma un valor particular (por ej. el paciente es fumador)
- **Negativa:** la variable puede tomar un subconjunto de valores, descartando los demás (por ej. la polución no es alta o la polución es baja o media)
- **Evidencia virtual:** existe incertidumbre sobre el valor de la variable (por ej. 80 % seguro que los rayos X son positivos)

# Árboles

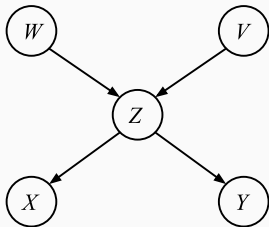
- Grafo donde existe un camino único entre cualquier par de nodos
- Para grafos dirigidos: cada nodo tiene un sólo padre



Tomada de PRML (Bishop 2009)

- Grafos dirigidos con nodos con más de un padre pero tienen un camino único entre cualquier par de nodos

Conexiones únicas



Conexiones múltiples

