Ejercicios 6: Redes Bayesianas

Departamento de Informática / Department of Computer Science
Universidad Carlos III de Madrid

Inteligencia Artificial
Grado en Ingeniería Informática
2019/20

Tabla de contenidos

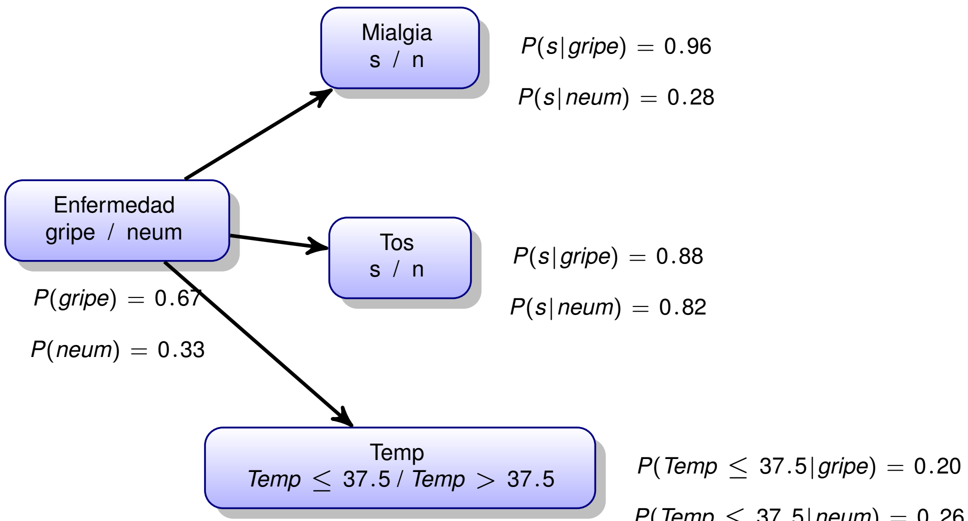
Notación y resumen

- Para representar un problema con una red Bayesiana hay que definir dos elementos, variables y dependencias:
  - Para representar las variables de la red tenemos que escribir su nombre y los valores que pueden tomar.
  - Para representar las relaciones entre variables podemos usar dos notaciones equivalentes:
    - Un diagrama donde las variables son los nodos y los arcos las dependencias, con las tablas de probabilidad condicional asociadas a cada nodo.
    - Simplemente el conjunto de tablas de probabilidad condicional.

  En ambos casos las variables hijo están condicionadas a sus variables padre ($P(hijo|\{padres\})$). Las variables que no tienen padre se etiquetan con sus probabilidades a priori.
- Para inferir algo en una RB nos preguntarán la probabilidad de que una variable tome un valor (consulta) sabiendo que ciertas variables tienen otros (evidencia) pero desconociendo los valores de otras variables (variables ocultas).
  - El método de inferencia exacta permite calcular numéricamente cualquier consulta.
  - En algunas ocasiones permitimos dejar los cálculos en función de un valor α de proporcionalidad (escalado); en los demás casos, hay que despejar α o calcular todos los valores posibles para la consulta en función de α y normalizar.

Ejercicio 1: Detección de gripe

Considere la red Bayesiana de la siguiente figura, que corresponde con un modelo de clasificador Naive Bayes:



Escriba la expresión que corresponde al cálculo que se pide a continuación, dejando sin calcular el factor de proporcionalidad α:

- Dada la evidencia $e = \{Temp > 37.5\}$, calcule la probabilidad de gripe (es decir, $P(gripe/Temp > 37.5)$)
- Dada la evidencia $e = \{Temp > 37.5, Tos = s\}$, calcule la probabilidad de gripe (es decir $P(gripe/Temp > 37.5, Tos = s)$)

Detección de gripe, $P(gripe/T > 37.5)$

$$P(E, T, D, Temp) = P(D/E)P(T/E)P(Temp/E)P(E)$$

$$P(E = gri/Temp > 37.5) = \alpha \sum_{d,t} P(D = d/E = gri)P(T = t/E = gri)P(Temp > 37.5/E = gri)P(E = gri) =$$

$$\alpha P(Temp > 37.5/E = gri)P(E = gri) \sum_{d} P(D = d/E = gri) \sum_{t} P(T = t/E = gri) =$$

$$\alpha P(Temp > 37.5/E = gri)P(E = gri) = \alpha 0.8 \times 0.67 = \alpha 0.536 \approx 0.7$$

$$P(E = neu/Temp > 37.5) = \alpha \sum_{d,t} P(D = d/E = neu)P(T = t/E = neu)P(Temp > 37.5/E = neu)P(E = neu) =$$

$$\alpha P(Temp > 37.5/E = neu)P(E = neu) \sum_{d} P(D = d/E = neu) \sum_{t} P(T = t/E = neu) =$$

$$\alpha P(Temp > 37.5/E = neu)P(E = neu) = \alpha 0.74 \times 0.33 = \alpha 0.2442 \approx 0.3$$

$$\alpha = \frac{1}{0.536 + 0.2442}$$

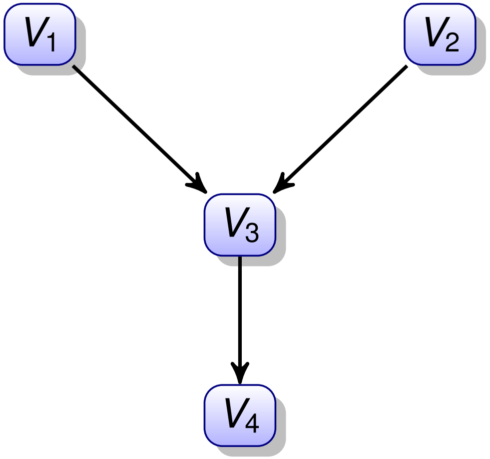
Detección de gripe, $P(gripe/T > 37.5, tos)$

$$P(E, T, D, Temp) = P(D/E)P(T/E)P(Temp/E)P(E)$$

$$P(gri/Temp > 37.5, tos) = \alpha P(tos/gri)P(Temp > 37.5/gri)P(gri)\sum_d P(D = d/E = gri) =$$
$$\alpha P(tos/gri)P(Temp > 37.5/gri)P(gri) = \alpha 0.88 \times 0.8 \times 0.67$$

Ejercicio 2: Inferencia, sólo fórmulas

Considere la red Bayesiana de la figura con cuatro variables binarias (booleanas):



1. Escriba cómo se puede escribir de forma simplificada la distribución de probabilidad conjunta para estas variables, a partir la estructura de ese grafo (factorización)
2. Escriba la expresión para la distribución de probabilidad marginal de P(V4)
3. Asuma que $V_2 = \mathrm{true}$. Cuál es la distribución de probabilidad *a posteriori* de V_2? Cuál es la distribución de probabilidad *a posteriori* de V_4?
4. Asuma que solo observamos que $V_4 = \textit{true}$ (es decir, V_2 es de nuevo desconocido). Calcule la distribución de probabilidad *a posteriori* de V_2.
5. Calcule la probabilidad conjunta de $V_1 = \textit{true}$, $V_2 = \textit{false}$, $V_3 = \textit{true}$, $V_4 = \textit{true}$.

Inferencia, sólo fórmulas, solución

1. $P(V_1, V_2, V_3, V_4) = P(V_4/V_3)P(V_3/V_1, V_2)P(V_1)P(V_2)$
2. $P(V_4) = \sum_{v1,v2,v3} P(V_4/V_3 = v3)P(V_3 = v3/V_1 = v1, V_2 = v2)P(V_1 = v1)P(V_2 = v2)$
3. $P(V_2/V_2 = true) = (1, 0)$
4. $P(V_2/V_4 = true)$

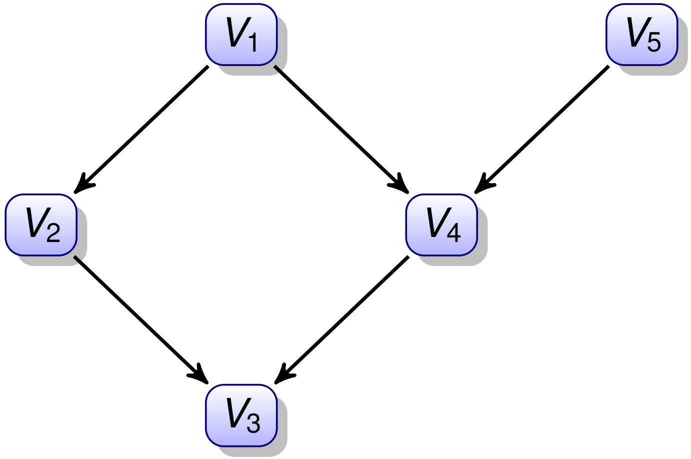
$$\begin{aligned}
&P(V_2 = true/V_4 = true) = \\
&= \alpha P(V_2 = true, V_4 = true) \\
&= \alpha \sum_{v1,v3} P(V_4 = true/V_3 = v3)P(V_3 = v3/V_1 = v1, V_2 = true)P(V_1 = v1)P(V_2 = true) \\
&= \alpha P(V_2 = true)\sum_{v1} P(V_1 = v1)\sum_{v3} P(V_4 = true/V_3 = v3)P(V_3 = v3/V_1 = v1, V_2 = true)
\end{aligned}$$

   $P(V_2 = false/V_4 = true)$ se calcula de forma similar. Después, podemos calcular $\alpha = \frac{1}{P(V_2=true/V_4=true)+P(V_2=false/V_4=true)}$

5. $P(V_1 = true, V_2 = false, V_3 = true, V_4 = true) = P(V_4 = true/V_3 = true)P(V_3 = true/V_1 = true, V_2 = false)P(V_1 = true)P(V_2 = false)$

Podéis inventar las probabilidades para esta red e intentar hacer los cálculos con números.

Dada la red Bayesiana con cinco variables booleanas de la figura:



$P(v_1)=0.2$	
$P(v_5)=0.8$	
$P(v_2\|v_1)=0.5$	$P(v_2\|\neg v_1)=0.4$
$P(v_3\|v_2, v_4)=0.4$	$P(v_3\|\neg v_2, v_4)=0.7$
$P(v_3\|v_2, \neg v_4)=0.3$	$P(v_3\|\neg v_2, \neg v_4)=0.6$
$P(v_4\|v_1, v_5)=0.6$	$P(v_4\|\neg v_1, v_5)=0.2$
$P(v_4\|v_1, \neg v_5)=0.0$	$P(v_4\|\neg v_1, \neg v_5)=1.0$

- Calcule la probabilidad $P(\neg v_5/v_3)$. Es decir, la probabilidad de que V_5 sea falso dado que V_3 es cierto.

Inferencia exacta, solución

$$
\begin{aligned}
P(\neg v_5|v_3) &= \alpha \sum_{v_1,v_2,v_4} P(v_3|v_2,v_4)P(v_2|v_1)P(v_4|v_1,\neg v_5)P(\neg v_5)P(v_1) \\
&= \alpha P(\neg v_5)\sum_{v_1} P(v_1)\sum_{v_2} P(v_2|v_1)\sum_{v_4} P(v_3|v_2,v_4)P(v_4|v_1,\neg v_5) \\
&= \alpha P(\neg v_5)\times
\end{aligned}
$$

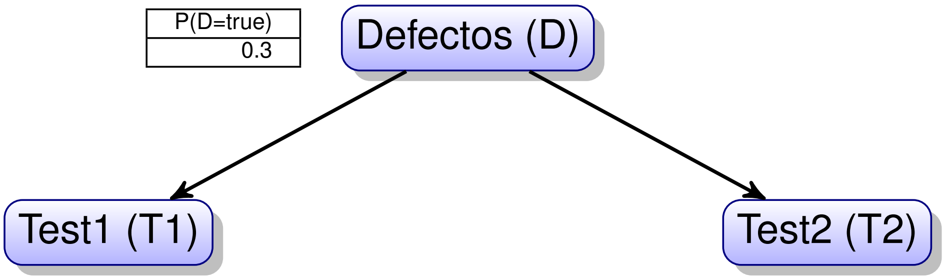
$$
\begin{aligned}
&[P(v_1) \\
&[P(v_2|v_1)[P(v_3|v_2,v_4)P(v_4|v_1,\neg v_5) + P(v_3|v_2,\neg v_4)P(\neg v_4|v_1,\neg v_5)]+ \\
&P(\neg v_2|v_1)[P(v_3|\neg v_2,v_4)P(v_4|v_1,\neg v_5) + P(v_3|\neg v_2,\neg v_4)P(\neg v_4|v_1,\neg v_5)]]+ \\
&P(\neg v_1) \\
&[P(v_2|\neg v_1)[P(v_3|v_2,v_4)P(v_4|\neg v_1,\neg v_5) + P(v_3|v_2,\neg v_4)P(\neg v_4|\neg v_1,\neg v_5)]+ \\
&P(\neg v_2|\neg v_1)[P(v_3|\neg v_2,v_4)P(v_4|\neg v_1,\neg v_5) + P(v_3|\neg v_2,\neg v_4)P(\neg v_4|\neg v_1,\neg v_5)]]]
\end{aligned}
$$

Todos estos valores son datos conocidos que se encuentran en las tablas de probabilidades condicionadas o bien son probabilidades a priori

Ejercicio 4: Concesionario

- Un vendedor de coches usados ofrece a los clientes potenciales que se realice una revisión en el coche que quieren comprar. El test debe revelar si el coche tiene o no defectos. La probabilidad a priori de que un coche tenga defectos es 0.3 %.
- Hay dos tests posibles:
  - **Test1** tiene tres posibles resultados: sin-defectos, defectos y no-conclusivo. Si el coche no tiene defectos, entonces las probabilidades para estos resultados son respectivamente 0.8, 0.05 y 0.15. Sin embargo, si tiene defectos, las probabilidades son 0.05, 0.75 y 0.2.
  - **Test2** tiene sólo dos posibles resultados: sin-defectos y defectos. Si el coche no tiene defectos, las probabilidades de cada resultado son 0.8 y 0.2 respectivamente. Y si los tiene, las probabilidades son 0.25 y 0.75.
- **Se pide:** Construir una red Bayesiana (estructura y TPC) que represente esta situación.

Concesionario, solución



P(D=true)
0.3

D	P(T1=SD)	P(T1=CD)	P(T1=NC)
T	0.05	0.75	0.2
F	0.8	0.05	0.15

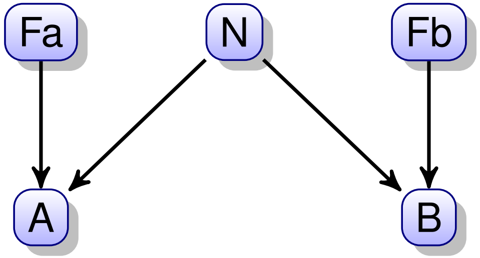
D	P(T2)=SD	P(T2)=CD
T	0.25	0.75
F	0.8	0.2

Ejercicio 5: Astrónomos

- Sean dos astrónomos (A and B) que miden el numero de estrellas que hay en una región del universo determinada, utilizando para ello sus respectivos telescopios, que tienen características similares.
  El 90 % de las veces la estimación que obtienen se corresponde con la real, mientras que en el 10 % restante la estimación varía en una estrella, siempre que los telescopios estén enfocados correctamente.
  Desgraciadamente, esto último solo ocurre en 9 de cada 10 ocasiones.
  Si no están enfocados correctamente, la estimación difiere del valor real en una estrella en un 50 % de los casos.
  Ambos saben que el número de estrellas en cada región seleccionada varía uniformemente entre 0 y 3, y la estimación debe también dar un valor entre 0 y 3.
  Usando esta información:
  1. Construya una red Bayesiana que represente el dominio
  2. Defina cada una de las tablas de probabilidad condicional de la red

Astrónomos, definiciones

- Variables:
  - *N*: Número de estrellas en la región del universo. $N \in \{0, 1, 2, 3\}$
  - *A*: Número de estrellas estimadas usando el telescopio A. $A \in \{0, 1, 2, 3\}$
  - *B*: Número de estrellas estimadas usando el telescopio B. $B \in \{0, 1, 2, 3\}$
  - *Fa*: El telescopio A está enfocado correctamente. $Fa \in \{f, \neg f\}$
  - *Fb*: El telescopio B está enfocado correctamente. $Fb \in \{f, \neg f\}$



Factorización de la probabilidad conjunta de esta red:

$$P(N, A, B, Fa, Fb) = P(A \mid Fa, N)\, P(B \mid N, Fb)\, P(Fa)\, P(Fb)\, P(N)$$

- $P(N=0) = P(N=1) = P(N=2) = P(N=3) = 0.25$
- $P(Fa = f) = 0.9$
- $P(Fa = \neg f) = 0.1$
- $P(Fb = f) = 0.9$
- $P(Fb = \neg f) = 0.1$

Astrónomos, Tablas de probabilidad condicional (TPC)

- $P(A \mid Fa, N)$

Fa	N	$P(A = 0 \mid Fa, N)$	$P(A = 1 \mid Fa, N)$	$P(A = 2 \mid Fa, N)$	$P(A = 3 \mid Fa, N)$
f	N=0	0.9	0.1	0	0
f	N=1	0.05	0.9	0.05	0
f	N=2	0	0.05	0.9	0.05
f	N=3	0	0	0.1	0.9
¬ f	N=0	0.5	0.5	0	0
¬ f	N=1	0.25	0.5	0.25	0
¬ f	N=2	0	0.25	0.5	0.25
¬ f	N=3	0	0	0.5	0.5

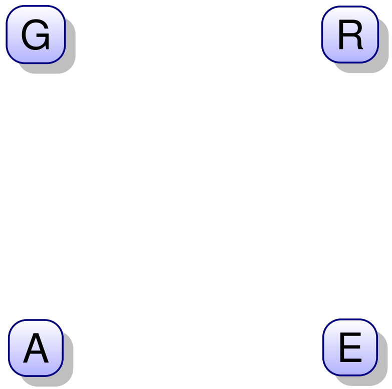
Ejercicio 6: Causas concurrentes

Cuando una persona alérgica a los gatos se encuentra en casa de un amigo y empieza a estornudar, puede ser por dos motivos: está resfriado o su amigo tiene un gato. Normalmente, los muebles de las casas con gatos están arañados.

1. Queremos construir un sistema capaz de razonar con la información del texto para personas alérgicas a los gatos. Diseña una red Bayesiana que represente esta información.
2. ¿Cúantos parámetros tiene la red?
3. Juan es alérgico a los gatos. Actualmente se encuentra en la casa de un amigo y observa que los muebles están arañados ¿Cuál es la expresión para calcular la probabilidad de que su amigo tenga un gato?
4. Después Juan comienza a estornudar ¿Cuál es ahora la expresión para calcular la probabilidad de que su amigo tenga un gato? ¿Cuál es la expresión para calcular la probabilidad de que Juan esté resfriado?

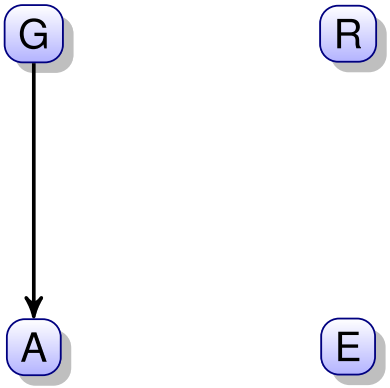
Nota: considere que el dominio considerado son las personas alérgicas a los gatos. ¿Cómo variaría el ejercicio si guera genérico (para todo tipo de personas), y qué otros datos podrían faltar en ese caso?

1. Se pueden identificar las siguientes variables aleatorias:
   - G. Los amigos de Juan tienen gato. $G \in \{g, \neg g\}$
   - R. Juan se ha resfriado.$R \in \{r, \neg r\}$
   - A. Los muebles están arañados. $A \in \{a, \neg a\}$
   - E. Juan estornuda $E \in \{e, \neg e\}$



- G. Los amigos de Juan tienen gato. $G \in \{g, \neg g\}$
- R. Juan se ha resfriado.$R \in \{r, \neg r\}$
- A. Los muebles están arañados. $A \in \{a, \neg a\}$
- E. Juan estornuda $E \in \{e, \neg e\}$

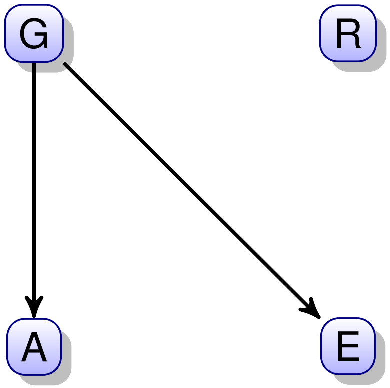
- La red bayesiana, usando las relaciones de causalidad, quedaría como sigue:



- G. Los amigos de Juan tienen gato. $G \in \{g, \neg g\}$
- R. Juan se ha resfriado.$R \in \{r, \neg r\}$
- A. Los muebles están arañados. $A \in \{a, \neg a\}$
- E. Juan estornuda $E \in \{e, \neg e\}$

- La red bayesiana, usando las relaciones de causalidad, quedaría como sigue:

  1. Tener gato (G) causa arañazos (A)

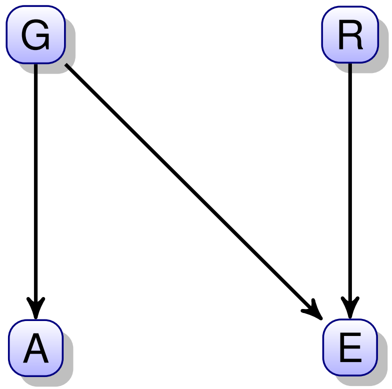


- G. Los amigos de Juan tienen gato. $G \in \{g, \neg g\}$
- R. Juan se ha resfriado.$R \in \{r, \neg r\}$
- A. Los muebles están arañados. $A \in \{a, \neg a\}$
- E. Juan estornuda $E \in \{e, \neg e\}$

- La red bayesiana, usando las relaciones de causalidad, quedaría como sigue:

    1. Tener gato (G) causa arañazos (A)
    2. Tener gato (G) causa estornudos (E)

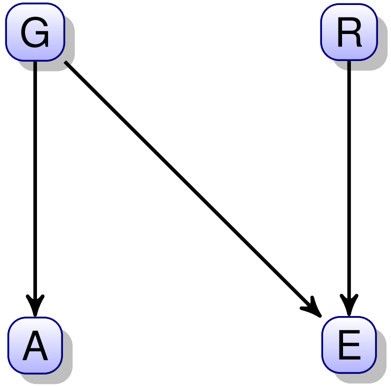
Causas concurrentes, Representación (2)



- G. Los amigos de Juan tienen gato. $G \in \{g, \neg g\}$
- R. Juan se ha resfriado.$R \in \{r, \neg r\}$
- A. Los muebles están arañados. $A \in \{a, \neg a\}$
- E. Juan estornuda $E \in \{e, \neg e\}$

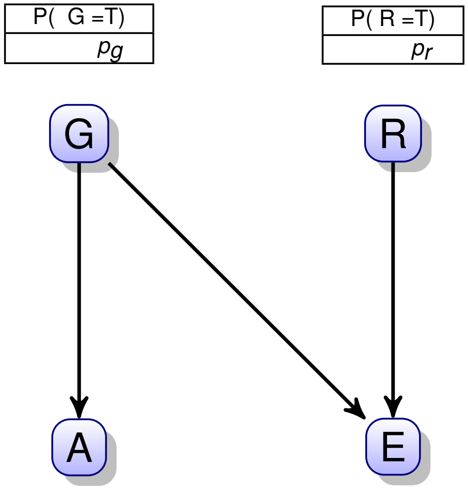
- La red bayesiana, usando las relaciones de causalidad, quedaría como sigue:

  1. Tener gato (G) causa arañazos (A)
  2. Tener gato (G) causa estornudos (E)
  3. Teber un resfriado (R) causa estornudos (E)



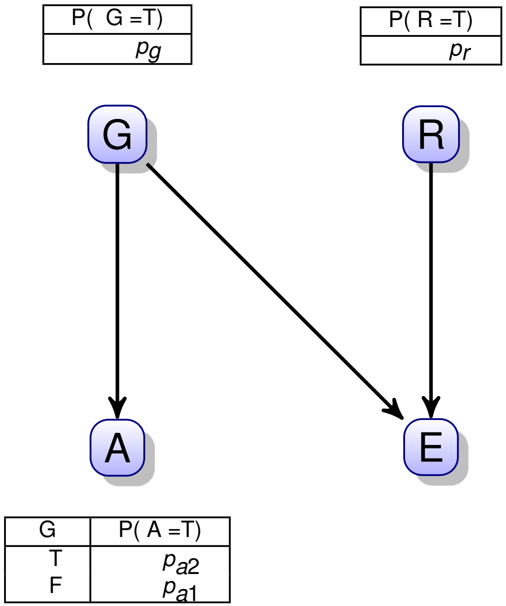
- $P(G)$ y $P(R)$ serán simplemente probabilidades a priori

Causas concurrentes, parámetros



- $P(G)$ y $P(R)$ serán simplemente probabilidades a priori
- A y E son condicionalmente independientes dado G

Causas concurrentes, parámetros

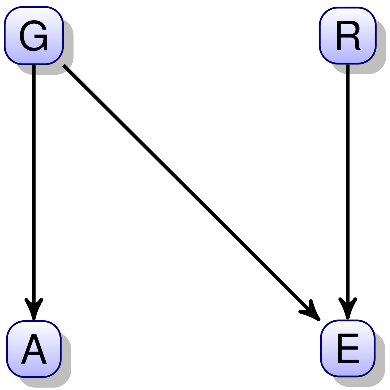


- $P(G)$ y $P(R)$ serán simplemente probabilidades a priori
- A y E son condicionalmente independientes dado G
- $P(A|G)$ es la probabilidad condicional asociada a la variable A

Causas concurrentes, parámetros

P(G =T)
p_g

P(R =T)
p_r



G	P(A =T)
T	p_{a2}
F	p_{a1}

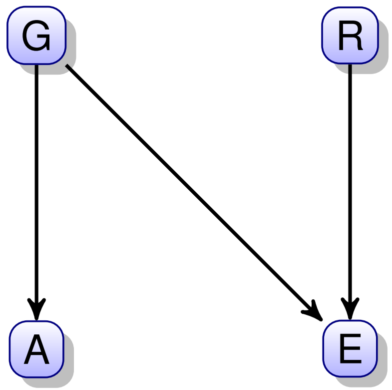
G	R	P(E =T)
T	T	p_{e11}
F	T	p_{e12}
T	F	p_{e21}
F	F	p_{e22}

- $P(G)$ y $P(R)$ serán simplemente probabilidades a priori
- A y E son condicionalmente independientes dado G
- $P(A|G)$ es la probabilidad condicional asociada a la variable A
- $P(E|G,R)$ es la probabilidad condicional asociada a la variable E

Causas concurrentes, parámetros

P(G =T)
p_g

P(R =T)
p_r



G	P(A =T)
T	p_{a2}
F	p_{a1}

G	R	P(E =T)
T	T	p_{e11}
F	T	p_{e12}
T	F	p_{e21}
F	F	p_{e22}

- $P(G)$ y $P(R)$ serán simplemente probabilidades a priori
- A y E son condicionalmente independientes dado G
- $P(A|G)$ es la probabilidad condicional asociada a la variable A
- $P(E|G,R)$ es la probabilidad condicional asociada a la variable E
- En la red, se cumple:

$$P(A,R,E,G) = P(G) \times P(R) \times P(A|G) \times P(E|G,R)$$

- En total se requieren 8 parámetros

Causas concurrentes, inferencia P(G =true / A =true)

$$
\begin{aligned}
P(G = true/A = true) &= \alpha P(G = true, A = true) \\
&= \alpha \sum_r \sum_e P(A = true, G = true, R = r, E = e)) \\
&= \alpha \sum_r \sum_e P(A = true/G = true)P(G = true)P(E = e/G = true, R = r)P(R = r) \\
&= \alpha P(A = true/G = true)P(G = true) \sum_r [P(R = r) \sum_e P(E = e/G = true, R = r)] \\
&= \alpha P(A = true/G = true)P(G = true) \sum_r [P(R = r) \sum_e P(E = e/G = true, R = r)] \\
&= \alpha P(A = true/G = true)P(G = true)
\end{aligned}
$$

Nota: se puede demostrar de forma general que las variables que no son ancestros de variables consulta o variables evidencia se acaban simplificando (son irrelevantes)

Causas concurrentes, inferencia P(G =true / A =true, E = true) y P(R =true / A =true, E = true)

$$
\begin{aligned}
P(G = t/A = t, E = t) &= \alpha P(G = t, A = t, E = t) \\
&= \alpha \sum_r P(A = t, G = t, R = r, E = t)) \\
&= \alpha \sum_r P(A = t/G = t)P(G = t)P(E = t/G = t, R = r)P(R = r) \\
&= \alpha P(A = t/G = t)P(G = t) \sum_r P(R = r)P(E = t/G = t, R = r)]
\end{aligned}
$$

$$
\begin{aligned}
P(R = t/A = t, E = t) &= \alpha P(R = t, A = t, E = t) \\
&= \alpha \sum_g P(A = t, G = g, R = t, E = t)) \\
&= \alpha \sum_g P(A = t/G = g)P(G = g)P(E = t/G = g, R = t)P(R = t) \\
&= \alpha P(R = t) \sum_g P(A = t/G = g)P(G = g)P(E = t/G = g, R = t)
\end{aligned}
$$

- Como siempre nos referíamos a personas alérgicas con la red diseñada es suficiente
- Si hubiera personas alérgicas y no alérgicas, habría una quinta variable booleana (Alérgico) que sería causa directa de la variable “Estornudar” (E).
- Habría que completar la TPC para E con los casos correspondientes a “No ser alérgico”, para lo que tendríamos que tener las probabilidades de estornudar si no se es alérgico, tanto en caso de que haya gato o no o de que se esté resfriado o no.
- Sin embargo, para el caso de Juan, el cálculo sería el mismo. Añadiríamos la evidencia “Alérgico=V” a la consulta, y en todas las expresiones aparecerían los mismos valores numéricos de las variables, los correspondientes a este caso.

Ejercicio 7: Detector de fraude

Imagina que trabajas en una institución financiera que necesita construir un aistema de detección de fraude.

- Cuando el titular de la tarjeta está de viaje al extranjero, son más probables las transacciones fraudulentas, ya que los turistas son un objetivo importante de los timadores. Concretamente, el 1 % de las transacciones que se realizan cuando el titular de la tarjeta está de viaje en el extranjero son fraudulentas, mientras que sólo el 0,2 % de la transacciones que se producen cuando no está viajando al extranjero son fraudulentas.
- En promedio, un 5 % de las transacciones suceden mientras el titular de la tarjeta está viajando al extranjero. Si una transacción es fraudulenta, entonces se incrementa la probabilidad de que la operación se deba a compras realizadas en el extranjero, a menos que el titular de la tarjeta esté viajando al extranjero. Concretamente, cuando el titular no está viajando al extranjero, el 10 % de las transacciones fraudulentas son compras realizadas en el extranjero, mientras que sólo un 1 % de las transacciones legítimas son compras realizadas en el extranjero.
- Por otro, cuando el titular está viajando al extranjero, el 90 % de las transacciones son compras realizadas en el extranjero, independientemente de la legitimidad de las mismas.

Ejercicio 7: Detector de fraude

Imagina que trabajas en una institución financiera que necesita construir un aistema de detección de fraude.

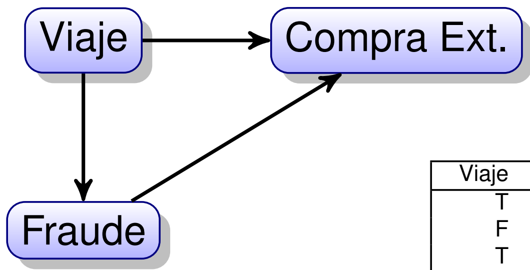
1. Construye una red bayesiana (estructura y distribuciones de probabilidad) para representar las relaciones del sistema de detección de fraudes.
2. Si el sistema ha detectado una compra fuera de la ciudad del titular. ¿Cuál es la probabilidad de fraude si no sabemos si el titular está de viaje o no?
3. Un agente llama al cliente para confirmar una transacción, pero el titular de la tarjeta no está en casa. Su esposa confirma que esta fuera de la ciudad en un viaje de negocios. ¿Cómo cambia la probabilidad de fraude considerando esta nueva información?

Detector de Fraude: representación

- Variables:
  - El titular está de viaje (Viaje): variable binaria
  - Hay fraude (Fraude): variable binaria
  - Hay compras realizadas fuera de la ciudad del titular (Compra Ext.): variable binaria
- Relaciones:
  - Inferencia causal: cuánto más alta es la probabilidad de estar de viaje, más alta es la probabilidad de fraude. Estar de viaje puede causar fraude.
  - Inferencia de diagnosis: cuánto más alta es la probabilidad de que haya una compra fuera de la ciudad, mayor es la probabilidad de fraude. Una compra fuera de la ciudad es una evidencia para fraude.
  - Inferencia intercausal: viajar y fraude pueden ser causa de una compra fuera de la ciudad. Estar de viaje explica las compras realizadas fuera de la ciudad y por lo tanto, es evidencia en contra de fraude.

Detector de fraude, estructura y TPC de la red

P(vi)
0.05



Viaje	Fraude	P(ce)
T	T	0.9
F	T	0.1
T	F	0.9
F	F	0.01

Viaje	P(fr)
T	0.01
F	0.002

Inferencia, P(Fraude =true| Compra Ext. =true)

- El sistema ha detectado una compra fuera de la ciudad del titular. ¿Cuál es la probabilidad de fraude si no sabemos si el titular está de viaje o no?

Inferencia, P(Fraude =true| Compra Ext. =true)

- El sistema ha detectado una compra fuera de la ciudad del titular. ¿Cuál es la probabilidad de fraude si no sabemos si el titular está de viaje o no?
- Consulta: Fraude
- Evidencia: Compra Ext.
- Variable oculta: Viaje
- Hay que calcular

Inferencia, P(Fraude =true| Compra Ext. =true)

- El sistema ha detectado una compra fuera de la ciudad del titular. ¿Cuál es la probabilidad de fraude si no sabemos si el titular está de viaje o no?
- Consulta: Fraude
- Evidencia: Compra Ext.
- Variable oculta: Viaje
- Hay que calcular
  P(Fraude =true| Compra Ext. =true) = P(fr | ce)

- P(fr | ce)
  - $= \alpha$ * [P(fr | vi) * P(ce | vi , fr) * P(vi)
    + P(fr |¬ vi) * P(ce |¬ vi , fr)*P(¬ vi)]
  - $= \alpha$ * [0.01 *0.90 *0.05 + 0.002* 0.10* 0.95]
  - $= \alpha$ * [0.00045 + 0.00019]
  - = 0.00064 α
- P(¬ fr | ce)
  - $= \alpha$ * [P(¬ fr | vi) * P(ce | vi ,¬ fr) * P(vi)
    + P(¬ fr |¬ vi) * P(ce |¬ vi ,¬ fr)*P(¬ vi)]
  - $= \alpha$ * [0.99 *0.90 *0.05 + 0.998* 0.01* 0.95]
  - $= \alpha$ * [0.04455+0.009481]
  - =0.054031 α
- α=1/(0.00064+ 0.054031)
- P(fr | ce)=1.1 %

Inference, P(Fraude = true| Compra Ext. =true, Viaje =true)

- Un agente llama al cliente para confirmar una transacción, pero el titular de la tarjeta no está en casa. Su esposa confirma que esta fuera de la ciudad en un viaje de negocios. ¿Cómo cambia la probabilidad de fraude considerando esta nueva información?

- P(fr | ce , vi)
  = α * [P(fr | vi) * P(ce | vi , fr) * P(vi)]
  = α * (0.01) (0.90) (0.05)
  = α * 0.00045
- P($\neg$ fr | ce , vi)
  = α * [P($\neg$ fr | vi) * P(ce | Viaje ,$\neg$ fr) * P(vi)]
  = α * (0.99) (0.90) (0.05)
  = α * * 0.04455
- α = 1/(0.00045+ 0.04455)
- P(fr | ce , vi) = 0.01

Ejercicio 8: Detector de incendios

Construya un clasificador Naive Bayes para la detección de fuegos a partir del siguiente conjunto de datos de entrenamiento. Los niveles de CO_2 y Temperatura se dan en rangos (entre 0 y 2).

1. Definir la estructura y tablas de probabilidad del clasificador
2. Determinar cuál sería la respuesta del clasificador para el caso marcado como "Test".
3. ¿Cómo se aplica el clasificador Naive Bayes para evaluar un caso de test en el que no se dispone del valor de algún atributo?

Humo	CO	CO_2	Temp	Fuego
S	N	1	2	S
S	S	2	0	S
N	S	2	1	S
N	N	1	2	S
S	S	1	1	S
S	S	0	2	S
N	N	2	2	S

Test

Humo	CO	CO_2	Temp	Fuego
S	N	1	1	?

Humo	CO	CO_2	T	Fuego
S	N	2	1	N
S	N	1	0	N
N	N	1	0	N
S	N	0	1	N
N	S	1	0	N
S	N	0	2	N
S	S	0	1	N

Test

Humo	CO	CO_2	T	Fuego
N	S	2	0	?

Detector de incendios, solución

Humo	CO	CO_2	Temp	Fuego
S	N	1	2	S
S	S	2	0	S
N	S	2	1	S
N	N	1	2	S
S	S	1	1	S
S	S	0	2	S
N	N	2	2	S
Test				
S	N	1	1	?

Humo	CO	CO_2	T	Fuego
S	N	2	1	N
S	N	1	0	N
N	N	1	0	N
S	N	0	1	N
N	S	1	0	N
S	N	0	2	N
S	S	0	1	N
Test				
N	S	2	0	?

- $P(F|Humo, CO, CO_2, T) = \alpha P(C)P(Humo|F)P(CO|F)P(CO_2|F)P(T|F)$
- La probabilidades a priori de la clase es: $P(F = S) = 0.5$
- Las probabilidades para la clase Fuego=S son:

$$P(H|F = S) = \langle \tfrac{4}{7}\ (S), \tfrac{3}{7}\ (N) \rangle$$
$$P(CO|F = S) = \langle \tfrac{4}{7}\ (S), \tfrac{3}{7}\ (N) \rangle$$
$$P(CO_2|F = S) = \langle \tfrac{1}{7}\ (0), \tfrac{3}{7}(1)\tfrac{3}{7}\ (2) \rangle$$
$$P(T|F = S) = \langle \tfrac{1}{7}\ (0), \tfrac{2}{7}\ (1)\tfrac{4}{7}\ (2) \rangle$$

- Las probabilidades para la clase Fuego=N son:

$$P(H|F = N) = \langle \tfrac{5}{7}\ (S), \tfrac{2}{7}\ (N) \rangle$$
$$P(CO|F = N) = \langle \tfrac{2}{7}\ (S), \tfrac{5}{7}\ (N) \rangle$$
$$P(CO_2|F = N) = \langle \tfrac{3}{7}\ (0), \tfrac{3}{7}\ (1)\tfrac{1}{7}\ (2) \rangle$$
$$P(T|F = N) = \langle \tfrac{3}{7}\ (0), \tfrac{3}{7}\ (1)\tfrac{1}{7}\ (2) \rangle$$

Detector de incendios, solución

Humo	CO	CO_2	Temp	Fuego
S	N	1	2	S
S	S	2	0	S
N	S	2	1	S
N	N	1	2	S
S	S	1	1	S
S	S	0	2	S
N	N	2	2	S
		Test		
S	N	1	1	?

Humo	CO	CO_2	T	Fuego
S	N	2	1	N
S	N	1	0	N
N	N	1	0	N
S	N	0	1	N
N	S	1	0	N
S	N	0	2	N
S	S	0	1	N
		Test		
N	S	2	0	?

- $P(F|Humo, CO, CO_2, T) = \alpha P(C)P(Humo|F)P(CO|F)P(CO_2|F)P(T|F)$
- Para el primer caso,

  $P(F = S|S, N, 1, 1) = \alpha 0.5 P(H = S|F = S)P(CO = N|F = S)P(CO_2 = 1|F = S)P(T = 1|F = S)$

  $P(F = S|S, N, 1, 1) = \alpha 0.5 \times \frac{4}{7} \times \frac{3}{7} \times \frac{3}{7} \times \frac{2}{7} = \alpha 0.5 \times \frac{72}{7^4}$

  $P(F = N|S, N, 1, 1) = \alpha 0.5 P(H = S|F = N)P(CO = N|F = N)P(CO_2 = 1|F = N)P(T = 1|F = N)$

  $P(F = N|S, N, 1, 1) = \alpha 0.5 \times \frac{5}{7} \times \frac{5}{7} \times \frac{3}{7} \times \frac{3}{7} = \alpha 0.5 \times \frac{225}{7^4}$
- Vemos que para este caso $P(F|S, N, 1, 1) \propto \langle 72, 225\rangle$, luego la clase más probable es Fuego=N.

Detector de incendios, solución

Humo	CO	CO_2	Temp	Fuego
S	N	1	2	S
S	S	2	0	S
N	S	2	1	S
N	N	1	2	S
S	S	1	1	S
S	S	0	2	S
N	N	2	2	S
Test				
S	N	1	1	?

Humo	CO	CO_2	T	Fuego
S	N	2	1	N
S	N	1	0	N
N	N	1	0	N
S	N	0	1	N
N	S	1	0	N
S	N	0	2	N
S	S	0	1	N
Test				
N	S	2	0	?

- $P(F|Humo, CO, CO_2, T) = \alpha P(C)P(Humo|F)P(CO|F)P(CO_2|F)P(T|F)$
- Para el segundo caso,

  $P(F = S|N, S, 2, 0) = \alpha 0.5 P(H = N|F = S)P(CO = S|F = S)P(CO_2 = 2|F = S)P(T = 0|F = S)$

  $P(F = S|N, S, 2, 0) = \alpha 0.5 \times \frac{3}{7} \times \frac{4}{7} \times \frac{3}{7} \times \frac{1}{7} = \alpha 0.5 \times \frac{72}{7^4}$

  $P(F = N|N, S, 2, 0) = \alpha 0.5 P(H = N|F = N)P(CO = S|F = N)P(CO_2 = 2|F = N)P(T = 0|F = N)$

  $P(F = N|N, S, 2, 0) = \alpha 0.5 \times \frac{2}{7} \times \frac{2}{7} \times \frac{1}{7} \times \frac{3}{7} = \alpha 0.5 \times \frac{12}{7^4}$
- Vemos que para este caso $P(F|N, S, 2, 0) \propto \langle 72, 12\rangle$, luego la clase más probable es Fuego=S.

Detector de incendios, solución

Humo	CO	CO_2	Temp	Fuego
S	N	1	2	S
S	S	2	0	S
N	S	2	1	S
N	N	1	2	S
S	S	1	1	S
S	S	0	2	S
N	N	2	2	S
Test				
S	N	1	1	?

Humo	CO	CO_2	T	Fuego
S	N	2	1	N
S	N	1	0	N
N	N	1	0	N
S	N	0	1	N
N	S	1	0	N
S	N	0	2	N
S	S	0	1	N
Test				
N	S	2	0	?

- $P(F|Humo, CO, CO_2, T) = \alpha P(C)P(Humo|F)P(CO|F)P(CO_2|F)P(T|F)$
- Si falta un atributo, la probabilidad conjunta sería la suma de probabilidades para cada posible valor de dicho atributo
- Como el atributo no es progenitor directo ni de la variable consulta ni de ningún otro atributos, se puede ignorar el término
- Por lo tanto, para cualquier conjunto de atributos disponible, podemos calcular la verosimilitud (probabilidad posterior) de cada clase multiplicando las probabilidades correspondientes de las TPC para cada atributo disponible, sin considerar los que faltan:

$$P(class = c|\{values \in A\}) \propto P(class = c) \times \prod_{i \in A} P(value_i|class = c)$$