Razonamiento con Incertidumbre

Lic. Carmen R. Garcia Perez

PATRONES DE INFERENCIA EN REDES BAYESIANAS

- Hay tres patrones importantes de inferencia en las redes bayesianas.

1. **Inferencia Causal (o Descendente)**

(de las causas a los efectos).

2. **Inferencia de Diagnóstico (Ascendente)**

(de los efectos a las causas).

3. **Inferencia Intercausal (Justificación)**

(entre las causas de un efecto común).

1. Inferencia Causal (o Descendente)

(de las causas a los efectos)

- Se desea calcular $p(S|O)$, la probabilidad de que el brazo se mueva dado que el Objeto se puede elevar.
- Como el Hecho de que el Objeto sea elevable es una de las causas de que el brazo pueda moverse, se dice que este cálculo es un ejemplo de **Razonamiento Causal**.
- *A O se lo denomina **evidencia utilizada** en la inferencia, y a S el **nodo de consulta**.*

1. Inferencia Causal (o Descendente)

- Así es como se realiza la inferencia: Primero se expande $p(S|O)$ en el sumatorio de dos probabilidades conjuntas (Porque se debe tener en cuenta al otro padre de S,B):

$$p(S|O)=p(S,B|O)+p(S,\neg B|O)$$

- Luego se utiliza la expresión de la regla encadenada, tomando en cuenta al otro padre.

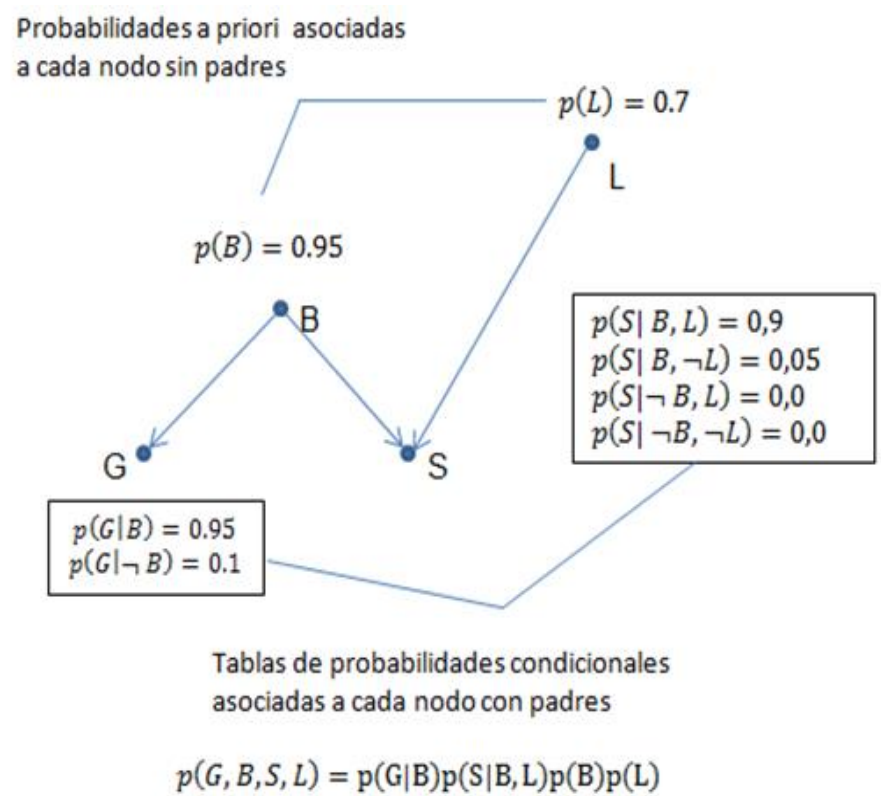
$$p(S|O)=p(S|B,O)p(B|O)+p(S|\neg B,O)p(\neg B|O)$$

1. Inferencia Causal (o Descendente)

- Pero $p(B|O)$=$p(B)$, fijémonos en que B no tiene padres e igualmente para, $p(\neg B|O)$=$p(\neg B)$.

- Por tanto, $p(S|O)$=$p(S|B,O)p(B) + p(S|\neg B,O)p(\neg B)$ como todas estas cantidades se dan en la red, se puede calcular y obtener:

$$p(S|O)=0{,}855$$



1. Inferencia Causal (o Descendente)

- Las operaciones que se han realizado pueden generalizarse a versiones más complejas de razonamiento causal. Las operaciones son:

  a) Escribir las probabilidades condicionales del nodo de consulta V, dada la evidencia en términos de la probabilidad conjunta de V y todos sus padres (los cuales no son la evidencia), dada la evidencia.

  b) Expresar esta probabilidad conjunta como la probabilidad de V, condicionada por todos sus padres.

2 . Inferencia de Diagnóstico (o Ascendente).

(de los efectos a las causas)

- Ahora se va a calcular $p(\neg O \mid \neg S)$, la probabilidad de que el bloque no se pueda elevar dado que el brazo no se mueve.
- Como se está utilizando un efecto(o síntoma) para inferir una causa, a este tipo de razonamiento se denomina ***razonamiento de diagnóstico.***

2 . Inferencia de Diagnóstico (o Ascendente).

$$p(\neg O|\neg S) = \frac{p(\neg S|\neg O)p(\neg O)}{p(\neg S)} (teorema\, de\, Bayes)$$

Ahora se calcula $p(\neg S|\neg O)$ = 0,9525 (mediante razonamiento causal), y luego se sustituye en:

$$p(\neg O|\neg S) = \frac{0{,}9525 \times 0{,}3}{p(\neg S)} = \frac{0{,}28575}{p(\neg S)}$$

De igual modo,

$$p(O|\neg S) = \frac{p(\neg S|O)p(O)}{p(\neg S)} = \frac{0{,}0597 \times 0{,}7}{p(\neg S)} = \frac{0{,}03665}{p(\neg S)}$$

Y como ambas expresiones deben sumar 1, se tiene $p(\neg O|\neg S)$ = 0,88632

2 . Inferencia de Diagnóstico (o Ascendente).

- Los cálculos que se ha realizado en este sencillo ejemplo de razonamiento de diagnóstico también se pueden generalizar.

- El paso principal es la utilización del teorema de Bayes para convertir el problema a uno de razonamiento causal.

3. Inferencia Intercausal (o Justificación).

(entre las causas de un efecto común)

- Si la única evidencia es que $\neg S$ (el brazo no se mueve), se puede calcular la probabilidad de que el bloque no se pueda elevar $\neg O$.
- Pero si también se tiene $\neg B$ (la batería no está cargada), entonces $\neg O$ debería ser menos cierto (menor peso).
- En este caso se dice que $\neg B$ *justifica* $\neg S$, haciendo a $\neg O$ menos cierto.
- Este tipo de inferencia utiliza un proceso de razonamiento causal(o descendente) incrustado en un proceso de diagnóstico(o ascendente).

3. Inferencia Intercausal (o Justificación).

$$p(\neg O|\neg B,\neg S) = \frac{p(\neg S,\neg B\,|\neg O)p(\neg O)}{p(\neg B,\neg S)} \quad \text{(Teorema de Bayes)}$$

$$= \frac{p(\neg S|\neg B,\neg O)p(\neg B|\neg O)p(\neg O)}{p(\neg B,\neg S)} \quad \text{(Def de la probabilidad condicional)}$$

$$= \frac{p(\neg S|\neg B,\neg O)p(\neg B)p(\neg O)}{p(\neg B,\neg S)} \quad \text{(Estructura de la red bayesiana)}$$

- Utilizando las probabilidades de la red, y resolviendo p($\neg B$, $\neg S$) de la forma habitual, se obtiene p($\neg O\,|\neg B$, $\neg S$)=0,030, es tal como se esperaba, mucho menor que p($\neg O\,|\neg S$). El uso del teorema de Bayes es un paso importante de la justificación.

EVIDENCIA CON INCERTIDUMBRE

- La expresión $p(V|\varepsilon)$, donde V es un nodo de consulta, no da la probabilidad correcta cuando la evidencia ε es incierta.
- En los cálculos sobre la red bayesiana, para poder obtener los nodos de evidencia, tener la certeza de la verdad o falsedad de las proposiciones que representan.
- Se puede alcanzar este requisito mediante una distribución en la que se tiene cada nodo de evidencia con un nodo hijo, del cual se tiene la certeza.

EVIDENCIA CON INCERTIDUMBRE

- Un claro ejemplo es la inferencia intercausal donde se puede suponer que el robot no tiene certeza acerca de si su brazo no se está moviendo(podría tener un sensor poco fidedigno).
- En este caso, la evidencia se podría obtener de un nodo S', que representaría la proposición ¨ el sensor del brazo dice que el brazo se ha movido¨. Se puede tener la certeza acerca de si dicha proposición es cierta o falsa dependiendo de su lectura.

EVIDENCIA CON INCERTIDUMBRE

- Entonces, la red bayesiana se utilizaria para calcular $p(\neg O | \neg B, \neg S')$ en vez de $p(\neg O | \neg B, \neg S)$. Desde luego, la red necesitaria los valores de $p(S' | S)$ y de $p(S' | \neg S)$ que representaría la fiabilidad del sensor.

Resumen

Definición de Redes Bayesianas:

Modelo gráfico probabilístico que permite describir distribuciones conjuntas complejas a partir de distribuciones Condicionales locales simples.

Es la forma más básica de obtener conocimiento de entornos inciertos(Incertidumbre).

El grafo dirigido Acíclico tiene Nodos (Variable aleatoria) y Arcos (Influencia directa de una variable sobre otra).

Resumen

INDEPENDENCIA CONDICIONAL

Una red bayesiana puede verse como una colección de sentencias que expresan independencia condicional.

- Si a y c son independientes:

  $$P(a|b,c)=P(a|b)$$

- Si a y b son independientes:

  $$P(a,b|c) = P(a|c)P(b|c)$$

Permite definir los procedimientos de Inferencia.

Resumen

INDEPENDENCIA

Permite simplificar los modelos probabilísticos aplicando la regla de la cadena:

$$P(a|b)=P(a)$$

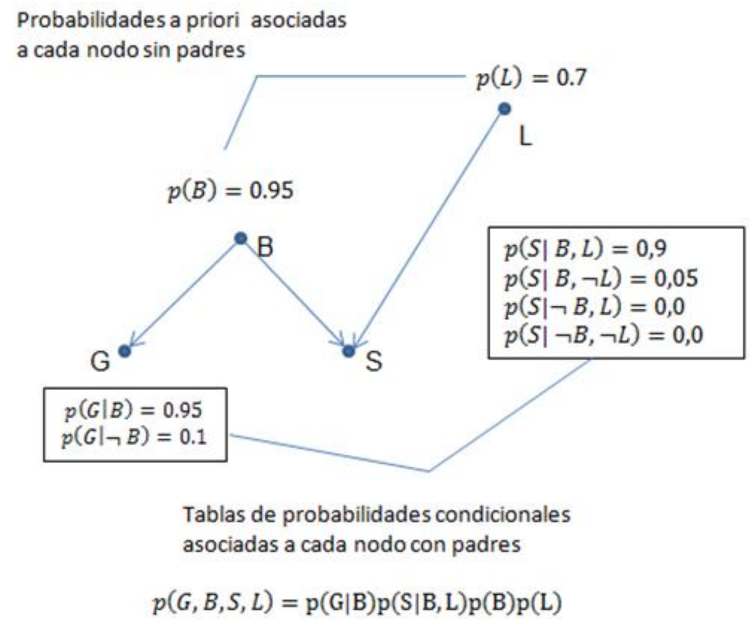
$$P(a,b)=P(a|b)P(b)=P(a)P(b)$$

Cuando a y b son independientes.

por lo que se reduce drásticamente la complejidad.

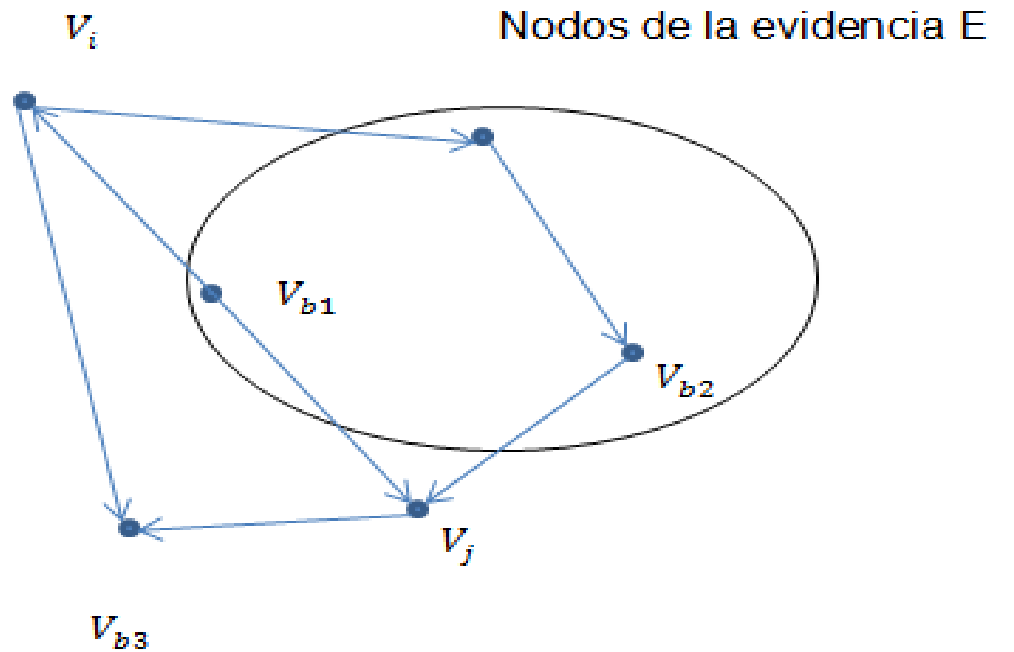
SEPARACIÓN- D

- Ocurre que una red bayesiana tiene más independencias condicionales que las que solo involucran a los padres de un nodo.
- Por ejemplo, $p(S|G,B)=p(S|B)$ es decir, S es condicionalmente independiente de *G* dado B.
  De forma intuitiva, el conocimiento del (efecto) *G* puede influir sobre el conocimiento de la (causa) B, el cual influye sobre el conocimiento del (otro efecto) S. Pero si dan la causa B, no hay nada que pueda decir *G* acerca de *S*. En este caso, decimos que B ***separa-d*** (separación dependiente de la dirección) a S de G.

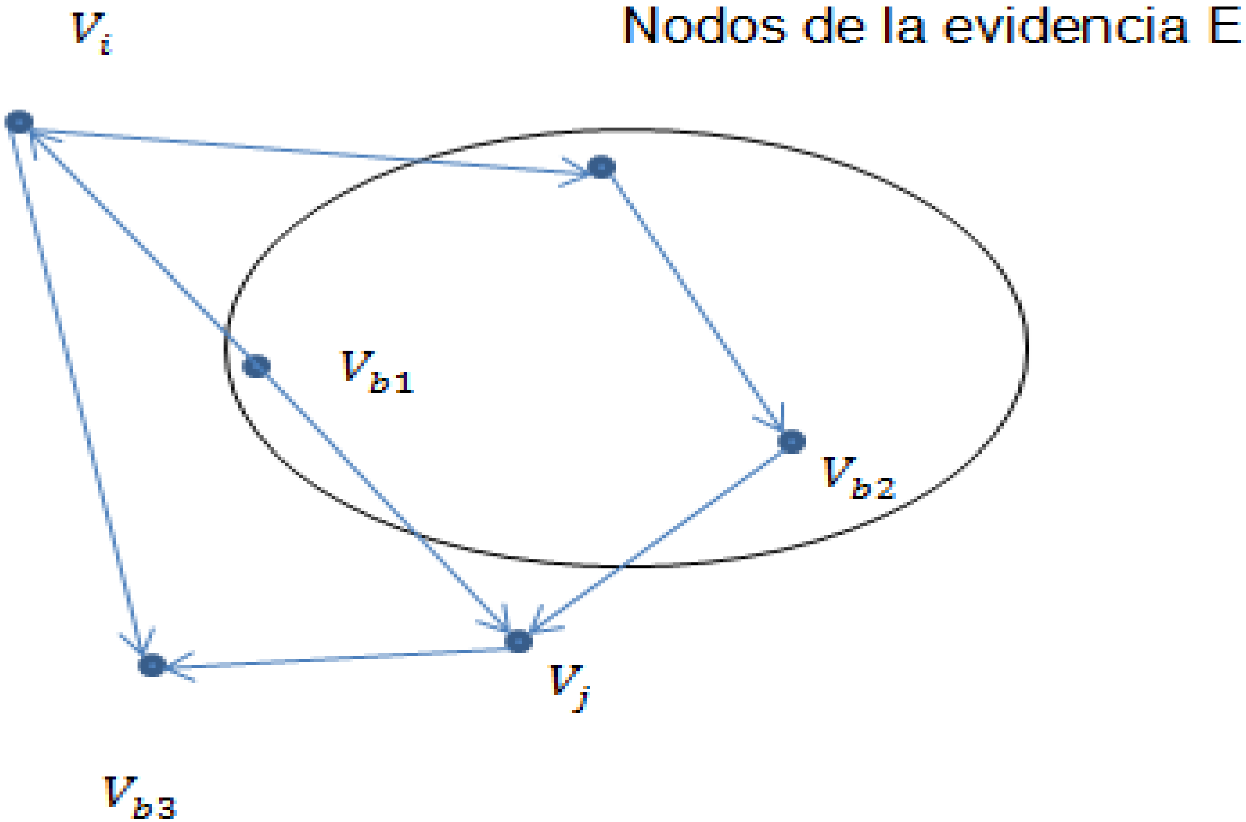


SEPARACIÓN- D

- Dos nodos Vi y Vj son independientes condicionalmente dado un conjunto de nodos ε (es decir, $I(Vi,Vj|\varepsilon)$) si por cada camino no dirigido entre Vi y Vj hay algún nodo V_b, que cumple algunas de las siguientes tres propiedades:
  1. V_b pertenece a ε, y ambos arcos salen de V_b,
  2. V_b pertenece a ε, y un arco va hacia V_b y el otro arco sale de él.
  3. Ni V_b ni ningún descendiente suyo pertenecen a ε, y ambos arcos van hacia V_b.



SEPARACIÓN- D



Dado el conjunto de nodos de la evidencia, V_i es independiente de V_j debido a que los tres caminos se bloquean entre sí. Los nodos que bloquean son:

- V_{b1} es un nodo evidencia, y ambos arcos salen de V_{b1}.
- V_{b2} es un nodo evidencia, y un arco va hacia él y otro sale de él.
- V_{b3} no es un nodo evidencia, ninguno de sus descendientes, y los dos arcos van a él.

SEPARACIÓN- D

- Cuando cualquiera de estas tres condiciones se cumple, se dice que el nodo V_b bloquea el camino dado ε.
- Tomar en cuenta que las rutas de la que se habla son caminos indirectos, es decir, caminos que ignoran la dirección de los arcos. Si todos los caminos entre V_i y V_j están bloqueados, entonces se dice que ε Separa-D V_i de V_j el cual concluimos que son independientemente condicionalmente dado ε.

SEPARACIÓN- D

Independencia condicional de tripletes:

- Cadena Causal: $\bigcirc \rightarrow \bigcirc \rightarrow \bigcirc$

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

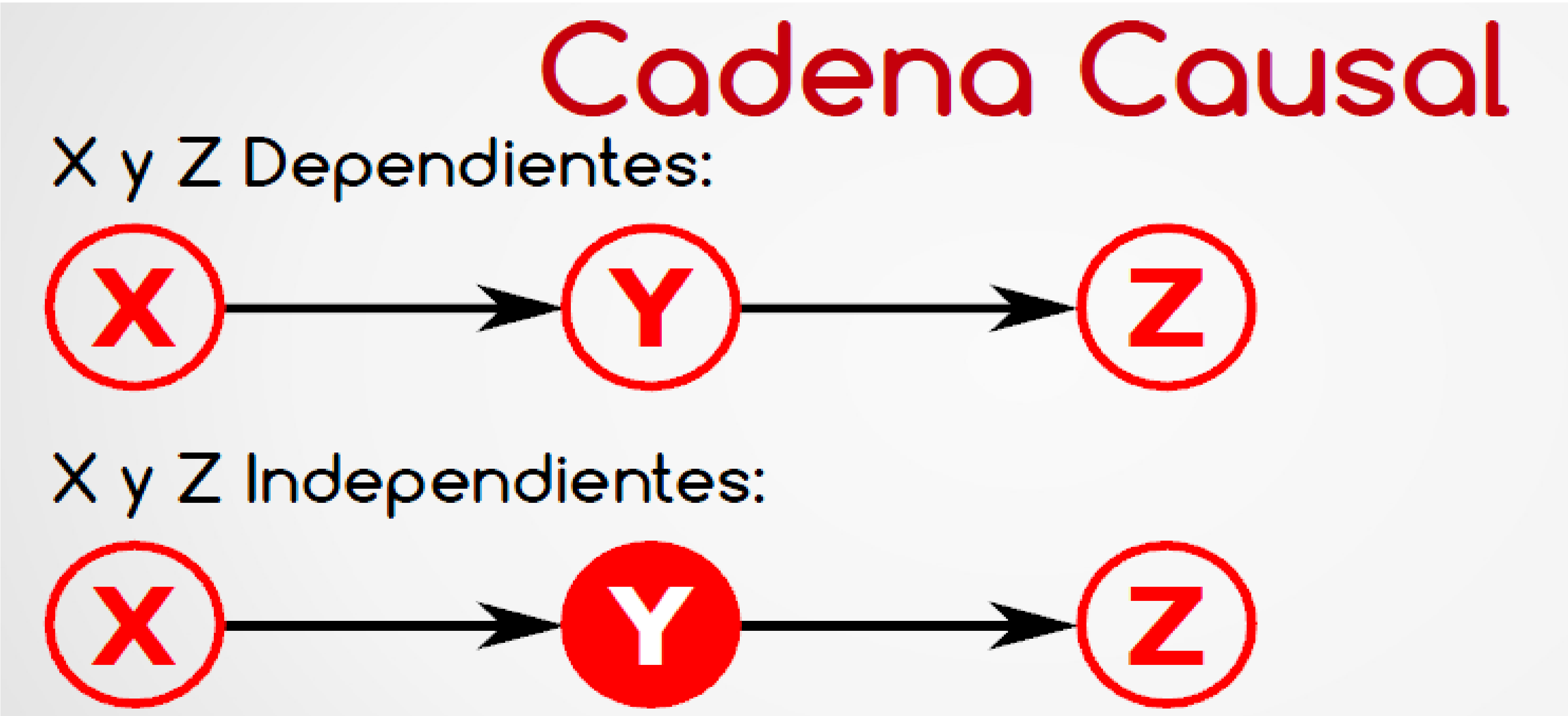
- Causa Común: $\bigcirc \leftarrow \bigcirc \rightarrow \bigcirc$

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

- Efecto Común: $\bigcirc \rightarrow \bigcirc \leftarrow \bigcirc$

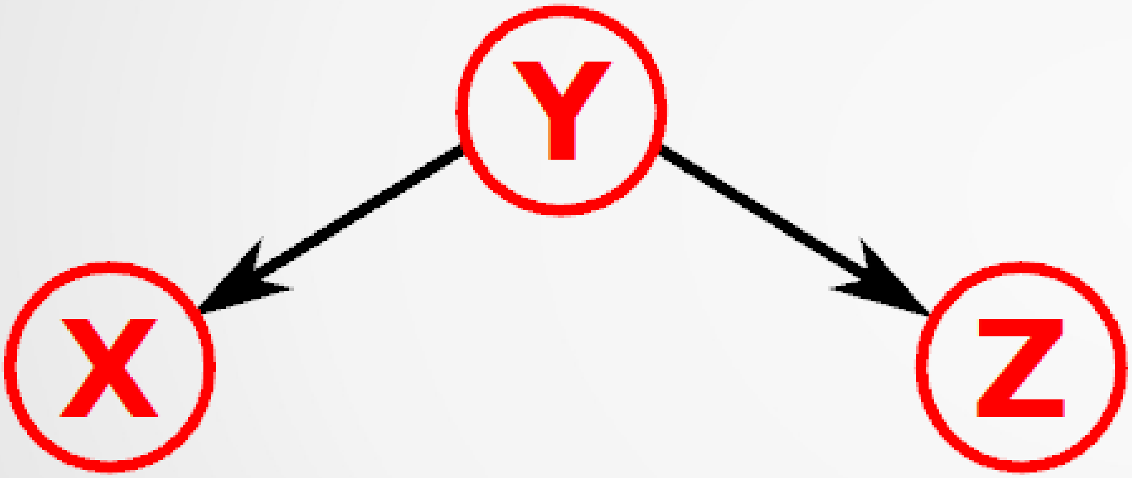
  $P(x,y,z)=P(x)\cdot P(z)\cdot P(y|x,z)$

  · Descencientes (de y)

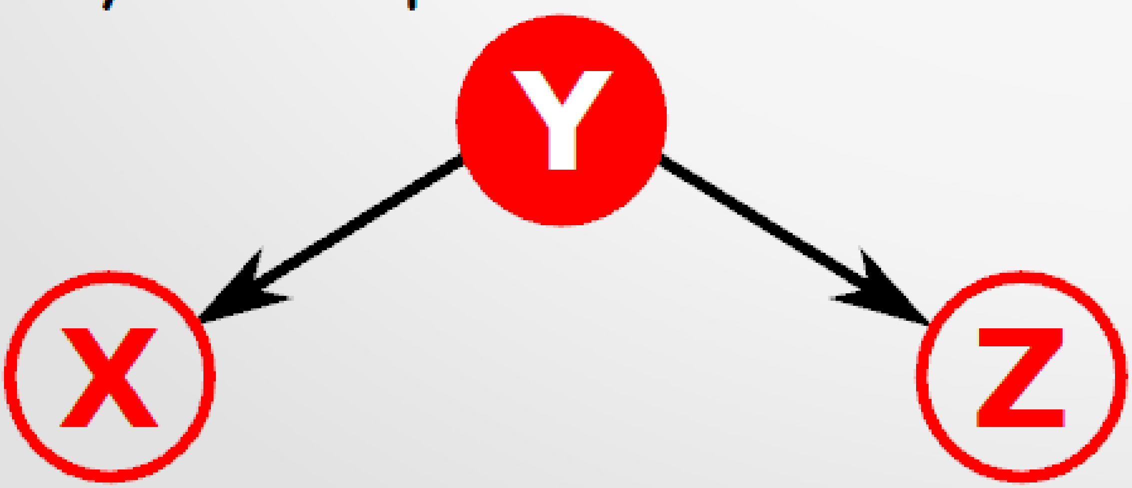

Cadena Causal
X y Z Dependientes:
X
Y
Z
X y Z Independientes:
X
Y
Z

Causa Común

X y Z Dependientes:

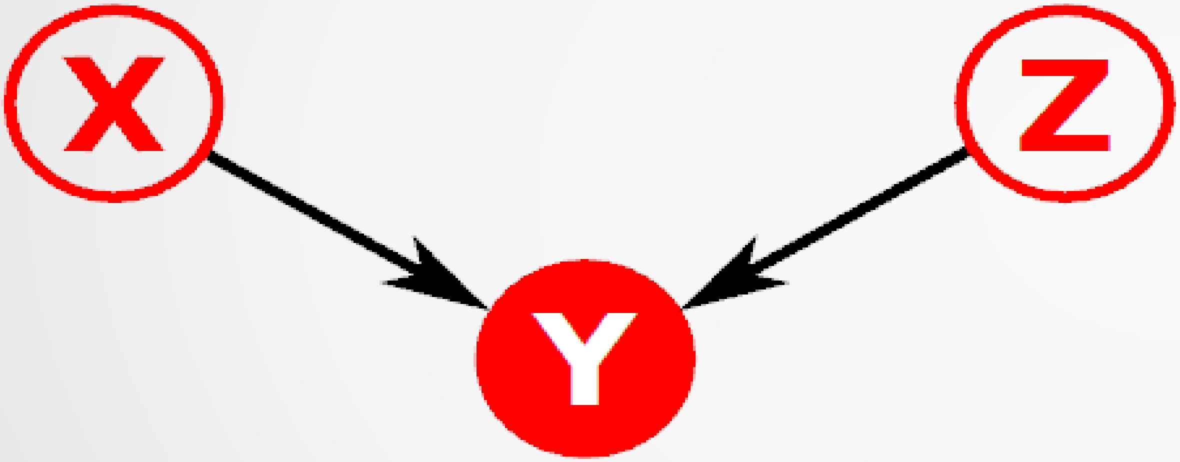


X y Z Independientes:



Efecto Común

X y Z Dependientes:



X y Z Independientes:

Efecto Común (Descendientes)

X y Z Dependientes: