Aprendizaje automatizado.

Tarea 4: Gráficos probabilísticos **_Flores Silva Pedro Arturo_**

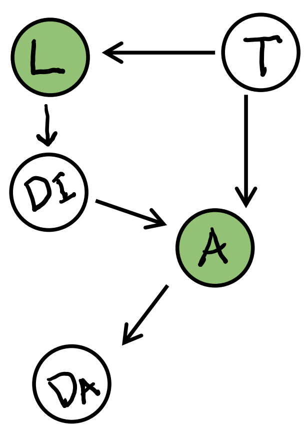
Ejercicio 1

En una planta nuclear hay una alarma que se activa cuando un indicador de temperatura excede un umbral. El indicador mide la temperatura del núcleo. Considera las variables booleanas A (alarma), D_A (alarma defectuosa), D_I (indicador defectuoso) y las variables enteras I (lectura del indicador) y T (temperatura real del núcleo).

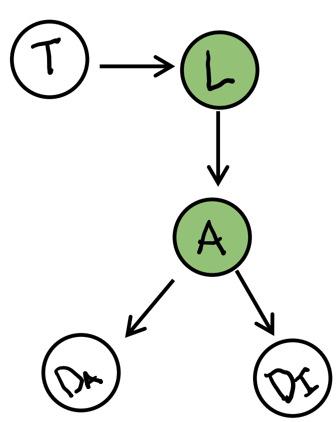
- Dibuja 3 redes bayesianas válidas diferentes que capturen el comportamiento del proceso, entre ellas incluye aquella que capture el mayor número de independencias condicionales con el menor número de arcos. Discute el modelo representado por cada una de las redes.
- De las redes dibujadas, escribe sus distribuciones conjuntas en términos de sus probabilidades condicionales.
- Asumiendo que las variables enteras I y T pueden tomar un máximo de 100 valores, ¿cuál sería el número de valores necesarios en cada nodo y el total en cada una de las redes?

Las tres redes se dibujan a continuación:

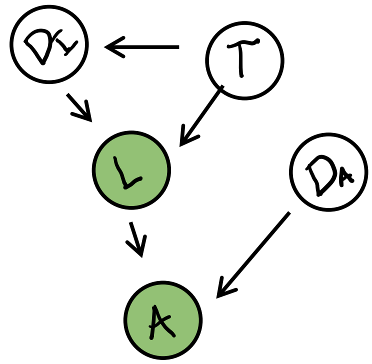
a)



b)



c)



Las tres redes bayesianas son válidas ya que son grafos acíclicos dirigidos. Antes de explicar el modelo de cada una de ellas debemos aclarar las variables.

A representa el estado de la alarma, activada (1) o apagada (0). ***T*** es la temperatura real del núcleo, que toma valores enteros. ***L*** es la lectura de la temperatura según el aparato indicador que es un aparato independiente de la alarma. ***DI*** es el estado del indicador, si el indicador es defectuoso asignamos 1, si no es defectuoso tiene valor cero. ***DA*** es el estado de la alarma, 1 si es defectuosa, 0 si no lo es.

A partir de aquí intuitivamente vemos que la temperatura del núcleo no va a depender de ninguna variable aquí considerada ya que es un valor físico. Es claro que la lectura del indicador va a depender de la temperatura y el hecho de que la alarma se active o no va a depender de la lectura del indicador. A partir de esta intuición se construyeron las tres redes bayesianas.

El modelo representado por cada una de ellas se explica a continuación:

a) Aquí consideramos el estado de la alarma que depende de la temperatura y de la lectura a través de la variable de si el detector es defectuoso o no, que a su vez éste depende de la lectura del mismo. Con este modelo podemos determinar si la alarma es defectuosa o no dependiendo de si se activó o no, que como ya se explico depende de la temperatura y la lectura.
b) Para este caso también nos interesa el estado de la alarma. Queremos saber la probabilidad de que esta se active dado la lectura del indicador. A partir de las probabilidades de activación podemos calcular las probabilidades de que tanto la alarma y el indicador sean defectuosos.
c) Para este caso modelamos la probabilidad de activación de la alarma basándonos en la probabilidad de que la alarma sea defectuosa y la lectura. Para el caso de la lectura, ésta depende de la temperatura y la probabilidad de que el instrumento falle dada una temperatura del núcleo. Esto es muy común en los aparatos electrónicos de medición. Que tienen un rango de temperaturas para los cuales se garantiza una probabilidad muy alta de operación correcta. Puede pasar que la temperatura del núcleo supere este rango del aparato y por ende falle, es decir, ¿Cuál es la probabilidad de que el aparato sea defectuoso dado un conjunto de temperaturas? Claramente la lectura depende de esta falla.

Las distribuciones conjuntas son:

a) $P(T, L, DI, A, DA) = p(T)p(L|T)p(DI|L)p(A|T, DI)p(DA|A)$

b) $P(T, L, DI, A, DA) = p(T)p(L|T)p(A|L)p(DI|A)p(DA|A)$

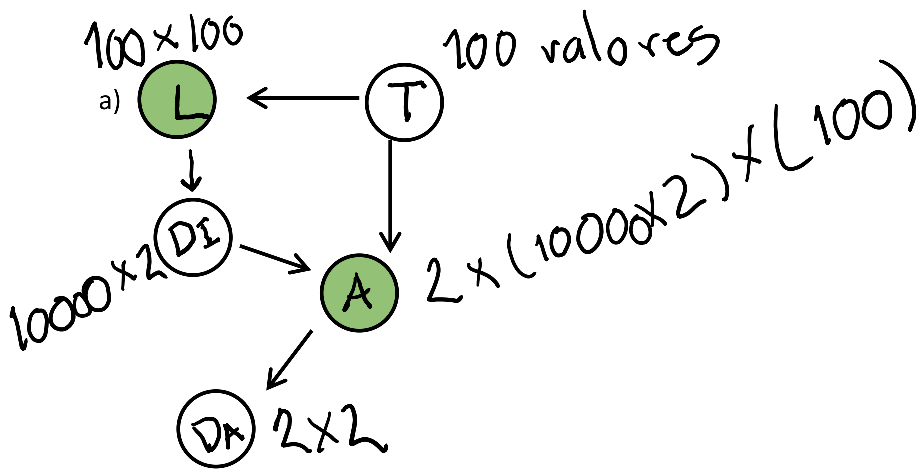
c) $P(T, L, DI, A, DA) = p(T)p(DI|T)p(L|DI, T)p(A|L, DA)$

Del diagrama los círculos pintados en verde son variables conocidas. A partir de ellas podemos inferir independencia condicional usando la d-separación, con el fin de encontrar la red que contenga menos independencias condicionales, que por simple inspección es la red b).

a) Dado que T es no conocido, no se asegura independencia condicional de $p(A|T, DI)$
b) En este caso no hay dependencias condicionales de más de una variable, por lo que no podemos reducir más la probabilidad conjunta
c) Dado que T y DI no se conocen no es posible asegurar la independencia condicional de $p(L|DI, T)$. Dado que se conoce A, L y DA no son condicionalmente independientes.

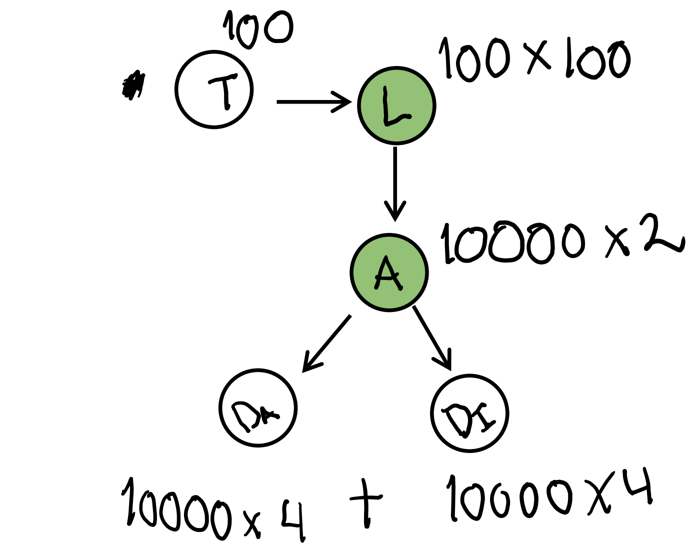
De esta forma la red que contiene menos dependencias condicionales es b) y también tiene la menor cantidad de arcos.

Para contar los valores necesarios en cada nodo volvemos a pintar las redes



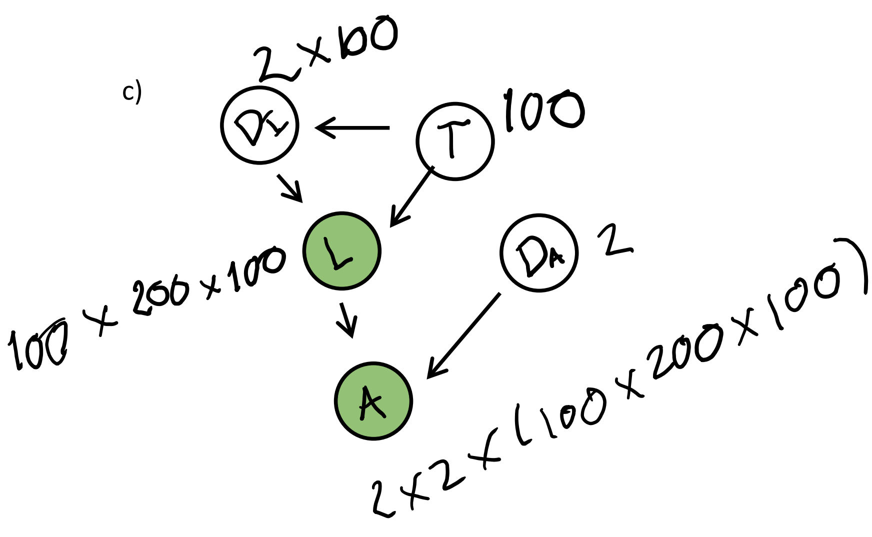
Dado que T puede tomar 100 valores entonces hay que evaluar para cada uno de estos valores 100 valores correspondientes a L, es decir para un valor fijo de T, hay 100 valores probables de L. Con esto para L tendrá 1000 (diez mil) valores. Dado que L tiene ahora diez mil valores, DI constará del cálculo de dos valores por cada valor fijo de L resultando en veinte mil (20000) valores. De forma análoga, para calcular los valores posibles de A dado T y DI, para un valor fijo de T y DI hay que calcular dos valores correspondientes a 0 o 1. Por lo que A tendrá 4 millones de valores. Finalmente Dado que hay que calcular dos valores por cada uno de los 4000000 valores de A, entonces al llegar a DA habríamos calculado 8000000 (8 millones) valores.

El total de valores es la suma de los valores



Nuevamente como T tiene cien valores, entonces para calcular los valores de L para un valor fijo de T se calculan 100 valores, por lo que tendremos 10000 diez mil valores, para cada valor fijo hay que calcular dos valores de cero y uno para el caso del estado de la alarma con lo cual tenemos 40000 cuarenta mil valores de probabilidad. Finalmente, para esos cuarenta mil hay que calcular otros dos para DA y otros dos para DI con esto tenemos que en total hay que calcular 80 mil probabilidades.

c)



Dado que T toma 100 valores y DI toma dos, entonces Para DI hay que calcular 200 valores. Para cada valor fijo de DI y T hay que calcular las probabilidades de cien lecturas, por lo que para L hay que calcular dos millones de valores. Finalmente, para calcular A hay que calcular para cada uno de los dos millones de valores 2 probabilidades multiplicado por 2 ya que también se considera el estado DA, por ende, tendríamos que haber calculado 8 millones de valores.

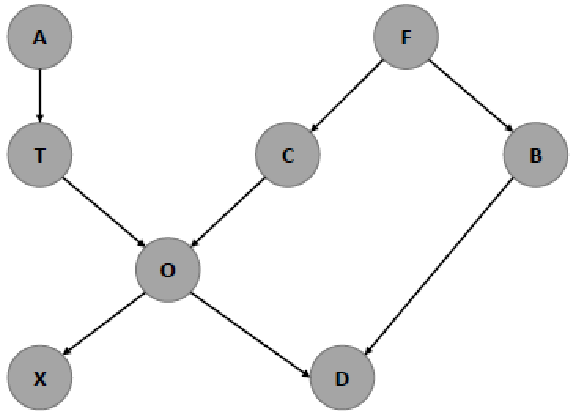


Figura 1: Modelo gráfico de diagnóstico de enfermedades de pulmón.

Ejercicio 2

A una clínica le concierne el diagnóstico de enfermedades de pulmón. Como se puede ver en el modelo de la Figura 1, una visita a Asia (A) hace que la probabilidad de tener tuberculosis (T) aumente. Los nodos en la gráfica tienen el siguiente significado:

Etiqueta	Significado
A	Visita a Asia
T	Tuberculosis
C	Cáncer de pulmón
O	Tuberculosis o Cáncer de pulmón
F	Fumador
X	Prueba de rayos X positiva
B	Bronquitis
D	Disnea

Di si las siguientes relaciones de independencia condicional son verdaderas o falsas y explica por qué:

a. $T \perp\!\!\!\perp F \mid D$

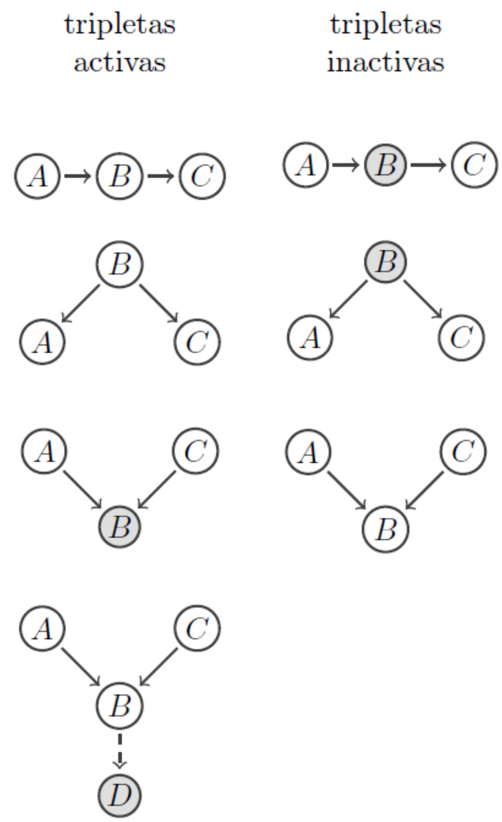
b. $C \perp\!\!\!\perp B \mid F$

c. $A \perp\!\!\!\perp F \mid C$

d. $A \perp\!\!\!\perp F \mid C, D$

El teorema de la d separacion nos dice que:

- X y Y son condicionalmente independientes dado Z si:
  - X y Y son separados por Z, es decir, no hay caminos activos entre X y Y.
- Un camino es activo si cada tripleta es activa.
  - cadena causal: $A \rightarrow B \rightarrow C$ donde B no ha sido observado,
  - causa común: $A \leftarrow B \rightarrow C$ donde B no ha sido observado,
  - efecto común: $A \rightarrow B \leftarrow C$ donde B o uno de sus descendientes D ha sido observado.



- $X \perp\!\!\!\perp Y \mid Z$?
  - Si hay un camino (no dirigido) activo entre X y Y entonces no se garantiza $X \perp\!\!\!\perp Y \mid Z$.
  - Si no hay caminos (no dirigidos) activos entre X y Y entonces se garantiza $X \perp\!\!\!\perp Y \mid Z$.

Bere et Richard

a) Del grafo, encontramos únicamente dos caminos no dirigidos del nodo T al nodo F y viceversa. Estos caminos son: $P(1) = \{T - O - C - F\}$ y $P(2) = \{T - O - D - B - F\}$. Para saber si T es condicionalmente independiente de F dado D debemos investigar estos dos caminos y encontrar que ambos están bloqueados (teorema de la d separación). Si existe alguno que no esté bloqueado, entonces no podemos asegurar la independencia condicional.
Para $P(1)$ tenemos las tripletas conectadas $\{T - O - C\}$ Y $\{O - C - F\}$ la primera de ellas es del tipo efecto común, dado que O no ha sido observado y tampoco sus descendientes, entonces esta tripleta es inactiva. La segunda tripleta es del tipo cadena causal, dado que ninguna de las tres variables ha sido observada entonces esta tripleta es activa, dado que la primera cadena bloquea este camino, entonces este camino no es activo.
Ahora para $P(2)$ tenemos las tripletas conectadas $\{T - O - D\}$, $\{O - D - B\}$ Y $\{D - B - F\}$, la primera tripleta es del tipo causal y es activa ya que O no ha sido observado. La segunda tripleta es del tipo efecto común y es activa ya que D ha sido observado. La tercera tripleta es del tipo causal y es activa ya que B no ha sido observado, de esta forma el camino es activo.
Dado que Existe al menos un camino activo ($P(2)$), entonces no podemos asegurar independencia condicional.

b) Del nodo C al B existen dos caminos: $P(1) = \{C - F - B\}$ Y $P(2) = \{C - O - D - B\}$.
Para $P(1)$ sólo existe una tripleta. Esta tripleta es del tipo causa común y es inactiva ya que F fue observado, por ende, este camino es inactivo.
Para $P(2)$ tenemos las tripletas conectadas $\{C - O - D\}$ Y $\{O - D - B\}$. La primera de ellas es una cadena causal y es activo ya que O no ha sido observado. La segunda de ellas es del tipo defecto común y es inactiva ya que D no ha sido observado, por lo que este camino es inactivo.
Dado que los dos caminos son inactivos, por el teorema de la d separación, podemos asegurar que C es condicionalmente independiente de B dado F.

c) Del nodo A al F existen dos caminos: $P(1) = \{A - T - O - C - F\}$ Y $P(2) = \{A - T - O - D - B - F\}$
Para $P(1)$ tenemos las tripletas $\{A - T - O\}, \{T - O - C\}$ Y $\{O - C - F\}$. La primera es del tipo cadena causal y es activa ya que no se ha observado T. La segunda tripleta es del tipo efecto común y es inactiva ya que no se ha observado O y sus descendientes no han sido observados, La última tripleta es del tipo cadena causal y es inactiva ya se ha observado C. Así dado que existe al menos una cadena inactiva que bloquea el camino, decimos que este camino es inactivo.
Para P(2) tenemos las siguientes tripletas: $\{A - T - O\}$, $\{T - O - D\}$, $\{O - D - B\}$ Y $\{D - B - F\}$. La primera tripleta es del tipo cadena causal y es activa ya que no se ha observado T. La segunda es también una cadena causal y es activa ya que no se ha observado O. La tercera tripleta es del tipo efecto común y es inactiva ya que no se ha observado D. Finalmente, la última tripleta es una cadena causal y es activa ya que no se ha observado B. Este camino es inactivo ya que la tercera tripleta es inactiva y bloquea el camino.
Dado que los únicos dos caminos disponibles son inactivos, entonces podemos asegurar la independencia condicional entre A y F dado C.

d) Del nodo A al F existen dos caminos: $P(1) = \{A - T - O - C - F\}$ Y $P(2) = \{A - T - O - D - B - F\}$
Para $P(1)$ tenemos las tripletas $\{A - T - O\}, \{T - O - C\}$ Y $\{O - C - F\}$. La primera es del tipo cadena causal y es activa ya que no se ha observado T. La segunda tripleta es del tipo efecto común y es activa ya que no se ha observado O pero su descendiente D sí ha sido observado, La última tripleta es del tipo cadena causal y es inactiva ya que se ha observado C. Así dado que existe una cadena inactiva que bloquea el camino, decimos que este camino es inactivo.
Para P(2) tenemos las siguientes tripletas: $\{A - T - O\}$, $\{T - O - D\}$, $\{O - D - B\}$ Y $\{D - B - F\}$. La primera tripleta es del tipo cadena causal y es activa ya que no se ha observado T. La segunda es también una cadena causal y es activa ya que no se ha observado O. La tercera tripleta es del tipo efecto común y es activa ya que se ha observado D. Finalmente, la última tripleta es una cadena causal y es activa ya que no se ha observado B. Este camino es activo ya que todas las tripletas son activas.
Dado que existe al menos un camino activo, entonces NO podemos asegurar la independencia condicional entre A y F dados C, D.

Ejercicio 3

Imagina una clínica que ayuda a pacientes con ébola en un área afectada por el virus. La red de la Figura 2 intenta capturar la dinámica por la cual las personas que sufren los síntomas pueden llegar a esta clínica y ver a un especialista. Existe la posibilidad de que alguien con ébola ($E = verdadero$) muestre síntomas, por ejemplo sangrado ($S = verdadero$) fiebre ($F = verdadero$) y visite la clínica ($V = verdadero$). El sangrado incrementa el riesgo de complicaciones ($C = verdadero$) y la persona puede ser llevada a ver un doctor especialista ($D = verdadero$).

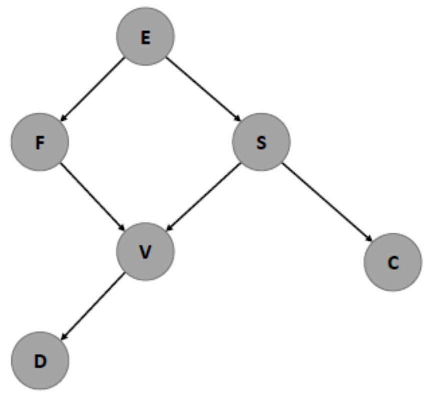


Figura 2: Modelo gráfico de detección de ébola.

- $P(E = verdadero) = 0.01$
- $P(F = verdadero|E = verdadero) = 0.6$
- $P(F = verdadero|E = falso) = 0.1$
- $P(S = verdadero|E = verdadero) = 0.8$
- $P(S = verdadero|E = falso) = 0.05$
- $P(V = verdadero|F = verdadero, S = verdadero) = 0.8$
- $P(V = verdadero|F = verdadero, S = falso) = 0.5$
- $P(V = verdadero|F = falso, S = verdadero) = 0.7$
- $P(V = verdadero|F = falso, S = falso) = 0.0$
- $P(C = verdadero|S = verdadero) = 0.75$
- $P(C = verdadero|S = falso) = 0.1$
- $P(D = verdadero|V = verdadero) = 0.6$
- $P(D = verdadero|V = falso) = 0.0$

Con estas probabilidades y el modelo gráfico de la Figura 2, realiza lo siguiente:

a. Escribe la distribución conjunta de la red bayesiana en función de las probabilidades condicionales.

b. Si un paciente es llevado al doctor ($D = verdadero$), usando un paquete de software calcula la probabilidad de que no tenga ébola ($P(E = falso|D = verdadero)$).

c. Convierte a la red bayesiana en un modelo gráfico no dirigido (campo aleatorio de Markov) y dibújalo. Captura tantas relaciones de independencia condicional como sea posible.

d. Debido a una campaña de concientización de la salud, las personas son alentadas a visitar la clínica en caso de que tengan fiebre. Esto incrementa la cantidad de visitas de personas con fiebre sin importar el estado de cualquier otra variable.

- ¿Qué probabilidades condicionales en la red se modifican debido a este cambio y en qué sentido?
- Describe cualquier efecto que esto tenga en la proporción de personas con complicaciones que visiten la clínica. Menciona exactamente qué probabilidades condicionales usaste para llegar a tu conclusión.

e. Asume que alguien que no tiene fiebre va al doctor, ¿qué relación de independencia condicional existe en la distribución que no puede ser descubierta a través del grafo solamente?

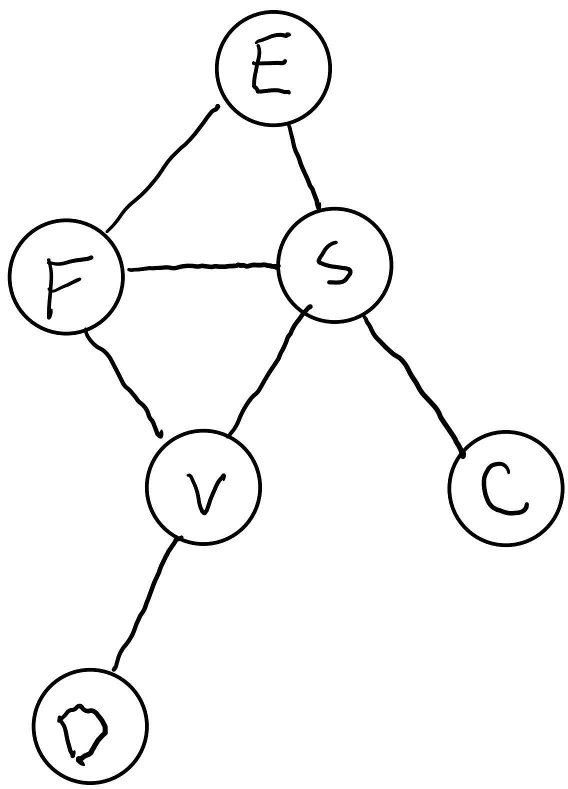
Opcional Calcula $P(V|E = verdadero)$ usando el método de eliminación de variables, describiendo de forma detallada el procedimiento que seguiste (2 puntos extras).

a) La distribución conjunta en función de las probabilidades condicionadas es:

$$P(E,F,S,V,C,D) = p(E)p(F|E)p(S|E)p(V|F,S)p(C|S)p(D|V)$$

b) Ver libreta redBayesiana.ipynb anexa en el correo.

c) El grafo que representa el campo aleatorio de Markov del grafo bayesiano se muestra a continuación:



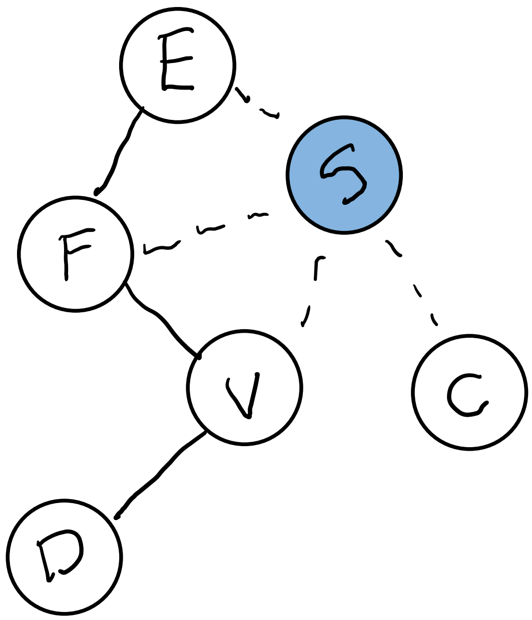
Para encontrar las relaciones de independencia condicional usamos las siguientes:

Definición: Sea $G = (V, A)$ un grafo no dirigido y X, Y, S conjuntos de vértices disjuntos de G, Entonces si cada camino entre un vértice en X y un vertice de Y contiene un vertice de S, entonces decimos que X, Y son u-separados por S

Proposición: Sea $G = (V, A)$ un grafo acíclico y dirigido con conjuntos disjuntos de vértices X, Y, S contenidos en G. Entonces, X, Y son d-separables por S sí y sólo sí X, Y son u-separados en el grafo no dirigido de G.

De esta forma podemos encontrar los siguientes subgrafos del campo aleatorio de markov:

1.



Si definimos el conjunto de vértices $X = \{E, F, V, D\}, S = \{S\} \;\; Y \; Y = \{C\}$ entonces X es u-separable de Y por S, por ende, X es d-separable de Y dado S por lo que tenemos las siguientes independencias condicionales:

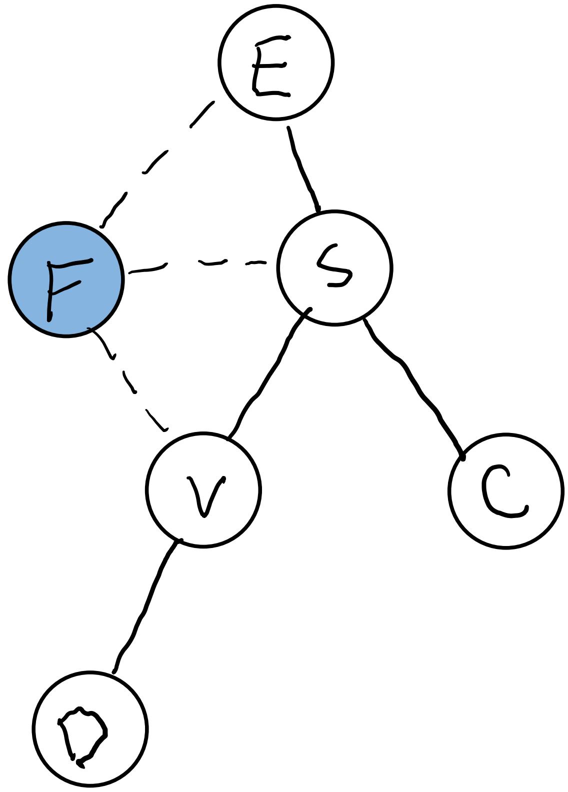
E es condicionalmente independiente de C dado S

F es condicionalmente independiente de C dado S

V es condicionalmente independiente de C dado S

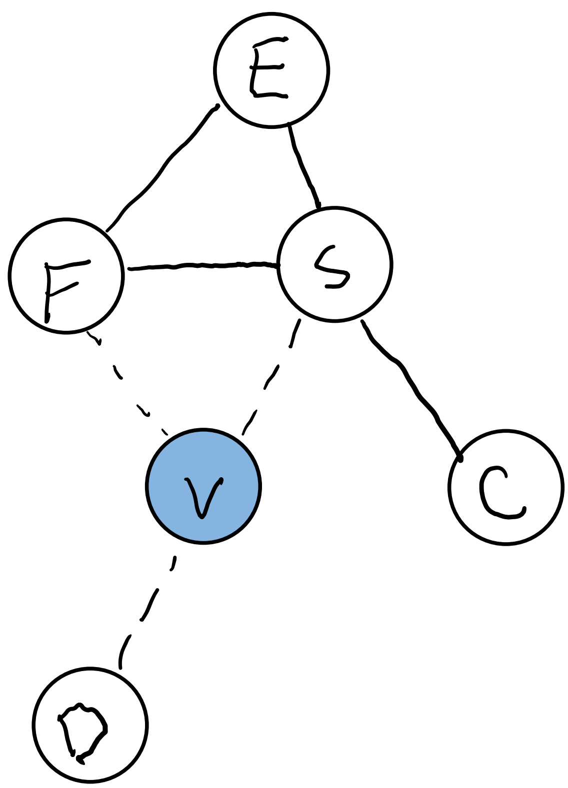
D es condicionalmente independiente de C dado S

2.



Para este caso no hay independencias condicionales ya que no hay conjuntos que puedan ser u-separables dado F

3.



Sea X = {E,F,S,C}, S = {V} y Y= {D}, entonces X, Y son u separables por S y por ende X, Y son d-separables dado V, con esto tenemos las siguientes independencias condicionales:

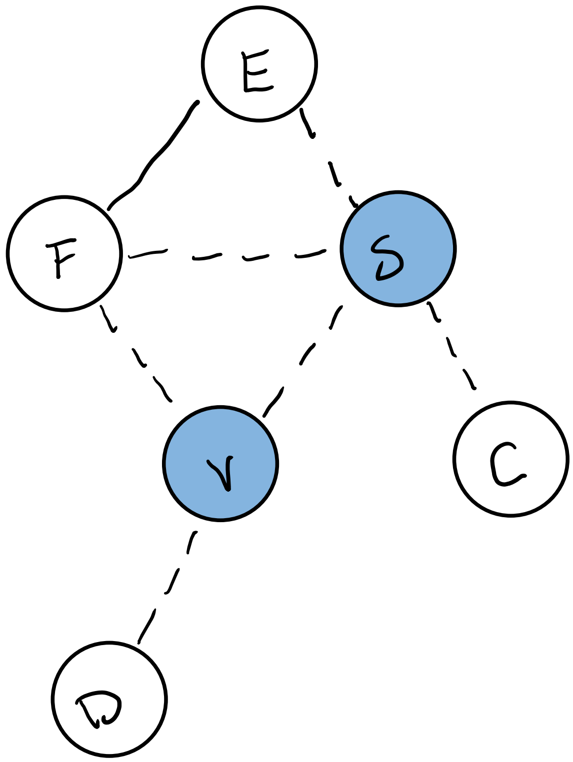
E es condicionalmente independiente de D dado V

F es condicionalmente independiente de D dado V

S es condicionalmente independiente de D dado V

C es condicionalmente independiente de D dado V

4.



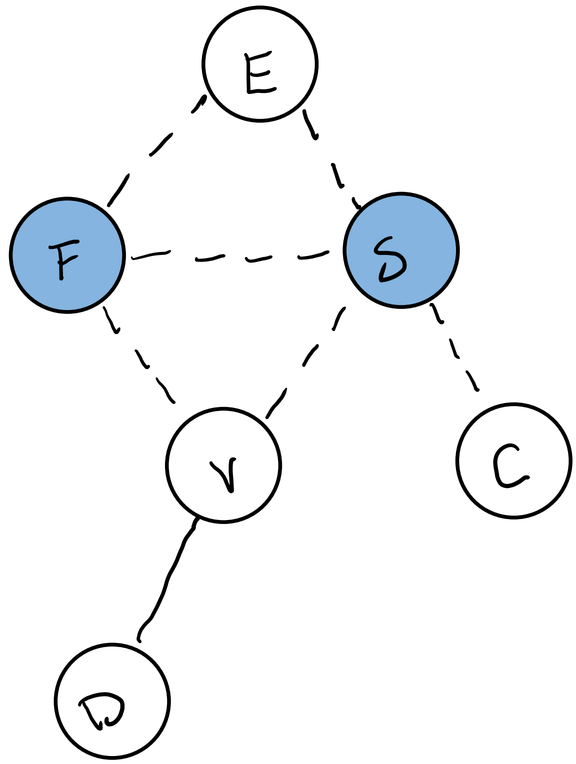
Sea X={E,F}, S={S,V}, Y={D,V}, entonces X, Y son u separables por S y por ende X, Y son d-separables dados S,V, con esto tenemos las siguientes independencias condicionales:

E es condicionalmente independiente de D y C dado V, S

F es condicionalmente independiente de D y C dados V, S

D es condicionalmente independiente de C dados V,S

5.



Sea X = {E}, S={F,S}, Y={V,C,D} entonces X, Y son u separables por S y por ende X, Y son d-separables dados F,S con esto tenemos las siguientes independencias condicionales:

E es condicionalmente independiente de V Dados F,S

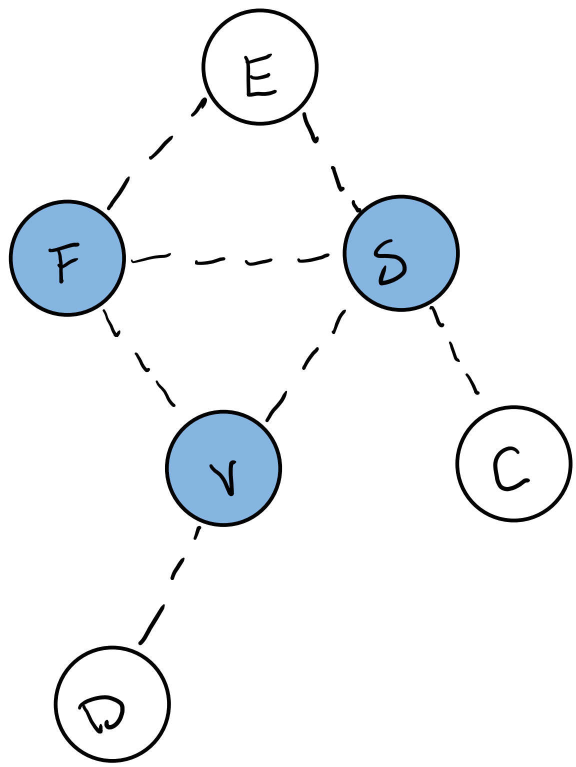
D es condicionalmente independiente de E dados F, S

C es condicionalmente independiente de E dados F, S

V es condicionalmente independiente de C dados F, S

D es condicionalmente independiente de C dados F,S

6.



Sea X={E}, S={F,S,V}, Y={D,C} entonces X, Y son u separables por S y por ende X, Y son d-separables dados F,S,V con esto tenemos las siguientes independencias condicionales:

E es condicionalmente independiente de C dados F, V, S

D es condicionalmente independiente de E dados F, V, S

D es condicionalmente independiente dado V

d)

1. Las probabilidades condicionales que cambian en la red debido a la campaña de concientización son:

   La probabilidad de visitar la clínica dado que se tiene fiebre debe aumentar ($P(V = verdadero \mid F = verdadero)$). Esto porque la indicación es que si tienes fiebre vayas a la clínica sin importar los otros síntomas.

   La probabilidad de no visitar la clínica dado que se tiene fiebre y se tenga o no sangrado ($P(V = verdadero \mid F = verdadero, S)$) deben aumentar, esto es por el caso anterior.

   La probabilidad de visitar la clínica dado que no se tiene fiebre, pero si se presenta sangrado ($P(V = verdadero \mid F = falso, S = verdadero)$) podría disminuir. Esto es un conocimiento a prior ya que la persona podría pensar que si no tiene fiebre entonces puede que no tenga ébola y por ende decida no ir a la clínica. Este conocimiento no necesariamente es verdad ya que depende fuertemente del espectador y el mensaje de la campaña, ya que puede pasar que el espectador piense que la fiebre es un factor más importante que el sangrado

2. Recordemos que las probabilidades de presentar complicaciones, según el modelo, depende de si se presenta sangrado o no, es decir $P(C = verdadero \,|S)$. Donde si se presenta sangrado la probabilidad de presentar complicaciones es muy alta (75%).
   Como ya se discutió la campaña de concientización aumenta la probabilidad de visitar la clínica dado que se tiene fiebre sin importar si se presenta sangrado o no. Dado que el modelo no relaciona directamente la fiebre con presentar complicaciones y/o la fiebre con el sangrado no podemos concluir nada acerca de la proporción de pacientes con complicaciones, y la proporción de pacientes con fiebre y sangrado y con fiebre sin sangrado se desconoce.
   Por otro lado, si asumimos que $P(V = verdadero \,|F = falso, S = verdadero)$ disminuye por la campaña de concientización, entonces claramente la proporción de pacientes con complicaciones disminuye considerablemente.
   En general para poder hacer una distinción en la proporción de complicaciones es necesario cambiar el modelo de tal forma que se relacione la complicación con la fiebre o el sangrado con la fiebre o en su caso cambiar la campaña de tal manera que se considere la fiebre y el sangrado.

e) El hecho de que una persona visite al doctor, es decir $D = verdadero$, sin presentar fiebre ($F = falso$) implica que estamos observando dichas variables. Sabemos que el hecho de D=verdadero, implica necesariamente que V = verdadero. Ahora bien, nos preguntamos sobre la independencia condicional de E con respecto a V dados D y F. Para ello notemos que existen dos caminos que llevan de E a V dados por: $P(1) = \{E - F - V\}$ Y $P(2) = \{E - S - V\}$. Usando el teorema de la d-separación vemos que el primer camino P(1) es inactivo ya que se trata de una cadena causal donde el nodo intermedio ha sido observado (F). Por otro lado el camino dos P(2) es activo ya que también se trata de una cadena causal pero el nodo intermedio S no ha sido observado, de esta forma el teorema no garantiza la independencia condicional entre E y V dados F y D.
Retomemos la definición de independencia condicional expresada matemáticamente: Sean a,b y c variables aleatorias, decimos que a es condicionalmente independiente de b dado c si se cumple que $p(a|b,c) = p(a|c)$, si consideramos la distribución conjunta de a y b condicionada en c podemos escribir equivalentemente la independencia condicional como: $p(a,b|c) = p(a|b,c)p(b|c) = p(a|c)p(b|c)$.
Entonces, dada la pregunta inicial, sin considerar el grafo, y la definición de independencia condicional nos preguntamos lo siguiente ¿ $\mathrm{p(E,V|F,D) \ = \ p(E|F,D)p(V|F,D)}$? que es la expresión matemática a la pregunta ¿Es E condicionalmente independiente de V dados F y D? Si esto es verdad basta con mostrar que la igualdad numérica es verdadera. Para ello usaremos el script desarrollado para el ejercicio 2 inciso b) (ver script) aquí solo se presentan los valores numéricos.
Parte derecha $\mathrm{p(E|F,D)p(V|F,D)}$:

$$\mathrm{p(E = verdadero|F = falso, D = verdadero) = 0.0670}$$
$$\mathrm{p(E = falso|F = falso, D = verdadero) = 0.9330}$$
$$\mathrm{p(V = verdadero|F = falso, D = verdadero) = 1.0000}$$
$$\mathrm{p(V = falso|F = falso, D = verdadero) = 0.0000}$$

Parte izquierda $\mathrm{p(E,V|F,D)}$:

$$\mathrm{p(E = verdadero, V = verdadero|F = falso, D = verdadero) = 0.0670}$$
$$\mathrm{p(E = verdadero, V = falso|F = falso, D = verdadero) = 0.0000}$$
$$\mathrm{p(E = falso, V = verdadero|F = falso, D = verdadero) = 0.9330}$$
$$\mathrm{p(E = falso, V = falso|F = falso, D = verdadero) = 0.0000}$$

De esta forma tenemos que:

$$0.0670 = \mathrm{p(E = verdadero, V = verdadero|F = falso, D = verdadero)}$$
$$= \mathrm{p(E = verdadero|F = falso, D = verdadero)}$$
$$* \mathrm{p(V = verdadero|F = falso, D = verdadero)} = 1.0000 * 0.0670 = 0.0670$$

Y

$$0.0000 = \mathrm{p(E = verdadero, V = falso|F = falso, D = verdadero)} = \mathrm{p(E = verdadero|F = falso, D = verdadero)} * \mathrm{p(V = falso|F = falso, D = verdadero)} = 0.0670 * 0.0000 = 0.0000$$

Y

$$0.9330 = \mathrm{p(E = falso, V = verdadero|F = falso, D = verdadero)}$$
$$= \mathrm{p(E = falso|F = falso, D = verdadero)} * \mathrm{p(V = verdadero|F = falso, D = verdadero)}$$
$$= 0.9330 * 1.0000 = 0.9330$$

Y

$$0.0000 = \mathrm{p(E = falso, V = falso|F = falso, D = verdadero)}$$
$$= \mathrm{p(E = falso|F = falso, D = verdadero)} * \mathrm{p(V = falso|F = falso, D = verdadero)}$$
$$= 0.9330 * 0.0000 = 0.0000$$

Así dado que la igualdad se cumple para todos los valores posibles de E y V dados F y D, decimos que E y V son condicionalmente independientes dados F y D.
Ya vimos que empleando el grafo y el teorema de la d-separación no fue posible asegurar la independencia condicional de E y V dados F y D, pero haciendo cálculos encontramos que realmente si son condicionalmente independientes.

Opcional:

De las probabilidades dadas podemos obtener las tablas de probabilidades condicionales:

Para el nodo E:

$P(E = verdadero)$= 0.01 $P(E = falso) = 0.99$

Para el nodo F:

$P(F = verdadero\ |E = verdadero) = 0.6$ $P(F = falso\ |E = verdadero) = 0.4$

$P(F = verdadero\ |E = falso) = 0.1$ $P(F = falso\ |E = falso) = 0.9$

Para el nodo S:

$P(S = verdadero\ |E = verdadero) = 0.8$ $P(S = falso\ |E = verdadero) = 0.2$

$P(S = verdadero\ |E = falso) = 0.05$ $P(S = falso\ |E = falso) = 0.95$

Para el nodo V:

$P(V = v\ |F = v,\ S = v) = 0.8$ $P(V = f\ |F = v,\ S = v) = 0.2$

$P(V = v\ |F = v,\ S = f) = 0.5$ $P(V = f\ |F = v,\ S = f) = 0.5$

$P(V = v \mid F = f,\ S = v) = 0.7$ $\qquad$ $P(V = f \mid F = f,\ S = v) = 0.3$

$P(V = v \mid F = f,\ S = f) = 0.00$ $\qquad$ $P(V = f \mid F = f,\ S = f) = 1.0$

Para el nodo C:

$P(C = verdadero \mid S = verdadero) = 0.75$ $\qquad$ $P(C = falso \mid S = verdadero) = 0.25$

$P(C = verdadero \mid S = falso) = 0.10$ $\qquad$ $P(C = falso \mid S = falso) = 0.90$

Para el nodo D:

$P(D = verdadero \mid V = verdadero) = 0.6$ $\qquad$ $P(D = falso \mid V = verdadero) = 0.4$

$P(D = verdadero \mid V = falso) = 0.0$ $\qquad$ $P(D = falso \mid V = falso) = 1.0$

Dado que queremos obtener $P(V \mid E = verdadero)$ y la distribución conjunta es: $P(E,F,S,V,C,D) = p(E)p(F|E)p(S|E)p(V|F,S)p(C|S)p(D|V)$, debemos marginalizar las variables ocultas (F,S,C,D). Entonces tenemos que la marginalización resulta en $P(V \mid E = verdadero) =$:

$$\sum_f \sum_s \sum_c \sum_d p(E)p(f|E)p(s|E)p(V|f,s)p(c|s)p(d|V)$$

Redistribuyendo las sumas tenemos que $P(V \mid E = verdadero) =$:

$$p(E)\left[\sum_d p(d|V)\right]\left[\sum_f \sum_s p(f|E)p(s|E)p(V|f,s)\sum_c p(c|s)\right]$$

Donde E va a tener valor verdadero.

Al eliminar d tenemos que:

$$f_d = \sum_d p(d|V)$$

$$f_d(V = verdadero) = 0.6 + 0.4 = 1.0$$

$$f_d(V = falso) = 0.0 + 1.0 = 1.0$$

Se sigue que:

$$p(V|E = v) = p(E)\left[\sum_f \sum_s p(f|E)p(s|E)p(V|f,s)\sum_c p(c|s)f_d(V)\right]$$

Ahora eliminando c se tiene que:

$$f_c(s) = \sum_c p(c|s)$$

$$f_c(S = verdadero) = 0.75 + 0.25 = 1.0$$

$$f_c(S = falso) = 0.1 + 0.9 = 1.0$$

Se sigue que:

$$p(V|E=v)=p(E)\left[\sum_f\sum_s p(f|E)p(s|E)p(V|f,s)f_c(s)f_d(V)\right]$$

Al eliminar s tenemos que

$$f_s(V,f)=\sum_s p(S|E=v)p(V|f,s)f_c(s)$$

S,V,F	$p(s\|E=v)p(V\|f,\ s)f_c(s)$
$verdadero, verdadero, verdadero$	$0.8*0.8*1.0=0.64$
$verdadero, verdadero, falso$	$0.8*0.7*1.0=0.56$
$verdadero, falso, verdadero$	$0.8*0.2*1.0=0.16$
$verdadero, falso, falso$	$0.8*0.3*1.0=0.24$
$falso, verdadero, verdadero$	$0.2*0.5*1.0=0.10$
$falso, verdadero, falso$	$0.2*0.0*1.0=0.00$
$falso, falso, verdadero$	$0.2*0.5*1.0=0.10$
$falso, falso\ falso$	$0.2*1.0*1.0=0.20$

Entonces:

$$f_s(V=v,f=v)=0.64+0.10=0.74$$

$$f_s(V=v,f=falso)=0.56+0.00=0.56$$

$$f_s(V=falso,f=v)=0.16+0.10=0.26$$

$$f_s(V=falso,f=falso)=0.24+0.20=0.44$$

Se sigue que:

$$p(V|E=v)=p(E)\left[\sum_f p(f|E)\,f_s(V,f)f_d(V)\right]$$

Al eliminar f tenemos que:

$$f_f(V)=\sum_f p(f|E)\,f_s(V,f)$$

F,V	$p(f\|E=v)f_s(V,f)$
$verdadero, verdadero$	$0.60*0.74=0.44$
$verdadero, falso$	$0.60*0.26=0.16$
$falso, verdadero$	$0.40*0.56=0.22$

$falso, falso$	$0.40 * 0.44 = 0.18$

Entonces:

$$f_f(V = verdadero) = 0.44 + 0.22 = 0.66$$

$$f_f(V = falso) = 0.16 + 0.18 = 0.34$$

Se sigue que:

$$p(V|E = v) = p(E)f_f(V)f_d(V)$$

Con esto ya podemos calcular $p(V|E = v)$ para V=verdadero y V=falso

V	$p(E)f_f(V)f_d(V)$
$verdadero$	$0.01 * 0.66 * 1.0 = 0.0066$
$falso$	$0.01 * 0.34 * 1.0 = 0.0034$

Así:

$$p(V = v|E = v) = \frac{0.0066}{0.0066 + 0.0034} = 0.66$$

$$p(V = f|E = v) = \frac{0.0034}{0.0066 + 0.0034} = 0.34$$