Práctica 4
Redes Bayesianas

Asignatura: Inteligencia Artificial
Fecha: 2/12/2020
Autor: Diego Marco Beisty, 755232

Universidad Zaragoza
1542

Hoja nº ....

Apellidos, Nombre: Marco Beisty, Diego

Titulación: .... Curso: .... Grupo: ....

Asignatura: .... Código: ....

Centro: ....

Fecha: .... /....... /.........

Calificación: ....

1

a) 1 $P(I \perp H \mid S)$ = Caminos (No cierto)

→ $\overbrace{IC}^{\text{Activo}}\underbrace{CP}_{\text{Activo}}\overbrace{PH}^{\text{Activo}}$ camino activo

→ $\overbrace{IC}^{\text{Activo}}\underbrace{CSH}_{\text{Inactivo}}$ camino inactivo

Como existe un camino activo las variables I y H son dependientes

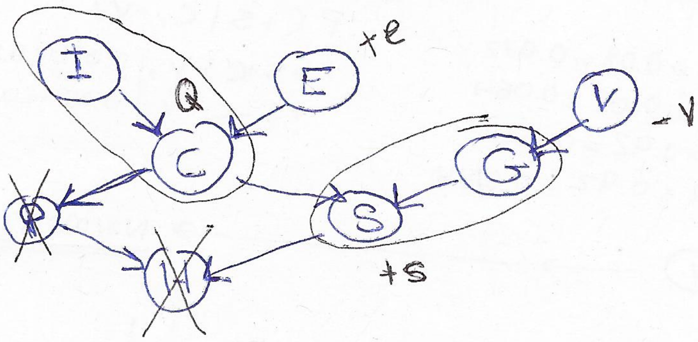
2 $P(V \perp H \mid S)$ = Caminos (No cierto)

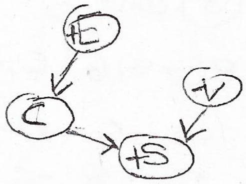
→ $\overbrace{VG}^{\text{Activo}}\underbrace{GSH}_{\text{Inactivo}}$ camino inactivo

→ $\overbrace{VG}^{\text{Act}}\overbrace{SC}^{\text{Activo}}$ $\underbrace{VGS}_{\text{Act}}\underbrace{CPH}_{\text{Act}}$ camino activo

Como sí existe un camino activo, las variables V y H son dependientes

b) $P(C \mid +e, +s, \neg v)$





Variable Query = C

Variables evidencia = E, S, V

Variables ocultas = I, G

1º P y H se pueden ignorar porque no son ancestros de ninguna variable query ni evidencia

2| Tablas de probabilidad con la evidencia.

I P(I)

+i	0.10
-i	0.90

E P(+e)

+e	0.05

C P(C|I,+e)

+e	+i	+c	0.01
+e	-i	+c	0.20
+e	+i	-c	0.99
+e	-i	-c	0.80

V P(-v)

-v	0.8

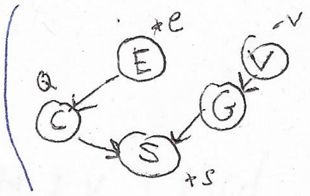
G P(G|-v)

-v	+G	0.08
-v	-G	0.92

S P(+s|C,G)

+c	+G	+s	0.90
+c	-G	+s	0.60
-c	+G	+s	0.80
-c	-G	+s	0.02

3| Junto I con C → Elimino I



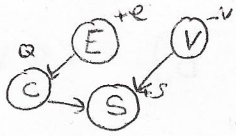
IC P(I,C)

+i	+c	0.01 × 0.10 = 0.001
-i	+c	0.20 × 0.90 = 0.18
+i	-c	0.99 × 0.10 = 0.099
-i	-c	0.80 × 0.90 = 0.72

P(C)

+c	0.001 + 0.18 = 0.181
-c	0.099 + 0.72 = 0.819

Junto G con S → Elimino G



P(+s,G|C,-v)

+s	+G	+c	-v	0.90 × 0.08 = 0.072
+s	+G	-c	-v	0.80 × 0.08 = 0.064
+s	-G	+c	-v	0.60 × 0.92 = 0.552
+s	-G	-c	-v	0.02 × 0.92 = 0.0184

P(+s|C,-v)

+s	+c	-v	0.072 + 0.552 = 0.624
+s	-c	-v	0.064 + 0.0184 = 0.0824

Junto E con C con S con V → Normalizo

P(+e, C, +s, -v)

+e	+c	+s	-v	0.05 × 0.181 × 0.624 × 0.8 = 0.00451
+e	-c	+s	-v	0.05 × 0.819 × 0.0824 × 0.8 = 0.00269

$$\frac{0.00451}{(0.00451+0.00269)} = 0.626$$

$$\frac{0.00269}{(0.00451+0.00269)} = 0.373$$

$$\boxed{P(C|+e,+s,-v) = 0.626}$$

Universidad Zaragoza 1542

Apellidos, Nombre: Marco Beisty, Diego

Titulación: .... Curso: .... Grupo: ....

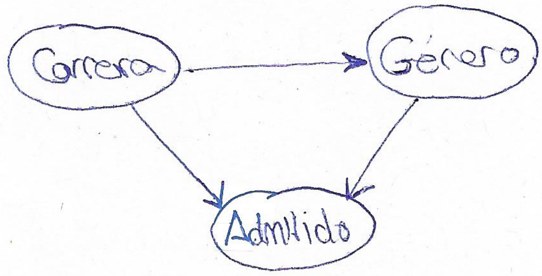
Asignatura: .... Código: ....

Centro: ....

Fecha: ...... /...... /........ Calificación: ....

2

a) variables aleatorias necesarias: Carrera, Género, Admitido



b)

C

	P(C)
A	$(825+108)/\overbrace{(825+108+560+25+325+593+417+375)}^{3228} = 0,289$
B	$(560+25)/3228 = 0,181$
C	$(325+593)/3228 = 0,284$
D	$(417+375)/3228 = 0.245$

G

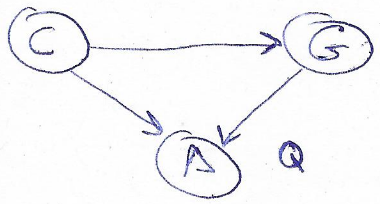
C	G	P(G|C)
A	M	$108/(108+825) = 0.115$
A	H	$825/(108+825) = 0,884$
B	M	$25/(25+560) = 0.0427$
B	H	$560/(25+560) = 0.957$
C	M	$593/(593+325) = 0.645$
C	H	$325/(593+325) = 0.354$
D	M	$375/(375+417) = 0.473$
D	H	$417/(417+375) = 0.526$

A

C	G	A	P(A|C, G)
A	M	Sí	$89/108 = 0.824$
A	H	Sí	$512/825 = 0.620$
B	M	Sí	$17/25 = 0.68$
B	H	Sí	$353/560 = 0.630$
C	M	Sí	$202/593 = 0.340$
C	H	Sí	$120/325 = 0.369$
D	M	Sí	$131/375 = 0.349$
D	H	Sí	$138/417 = 0.330$
A	M	No	0.176
A	H	No	0.38
B	M	No	0.32
B	H	No	0.37
C	M	No	0.66
C	H	No	0.631
D	M	No	0.651
D	H	No	0.67

c) Dada la tabla de probabilidad de A (admitidos), P(A | C, G)
La probabilidad de ser admitido en las carreras A, B y D siendo mujer es superior a la de ser admitido siendo hombre.
Sólo en la carrera C la probabilidad de ser admitido siendo hombre es superior a la de ser admitido siendo mujer.
Por lo tanto sí que ha podido haber discriminación de género hacia el hombre en la admisión.

D)



Q

~~No~~
Variable query = {A}.
Variables evidencia = {}
Variables ocultas = {C, G}

No se pueden ignorar las variables ocultas C y G puesto que son antecesores de la query A.

Junto C con G y con A ⟶ Elimino C

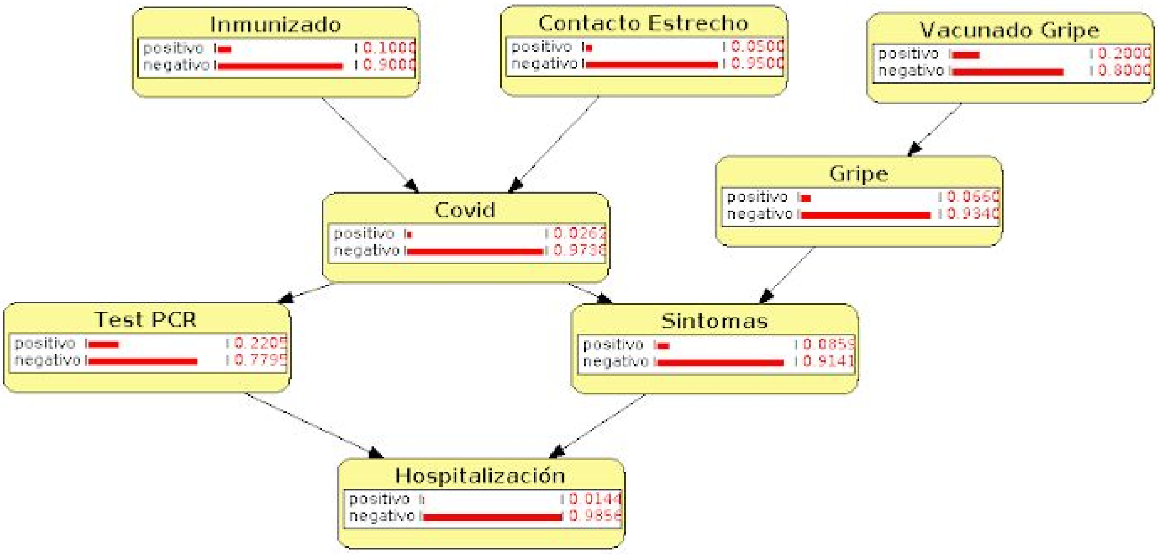
C	G	A	P(C, G, A)
A	M	Sí	0.824 x 0.115 x 0.289 = 0.027
A	H	Sí	0.620 x 0.884 x 0.289 = 0.158
B	M	Sí	0.680 x 0.0427 x 0.181 = 0.173
B	H	Sí	0.630 x 0.957 x 0.181 = 0.109
C	M	Sí	0.340 x 0.645 x 0.284 = 0.062
C	H	Sí	0.369 x 0.354 x 0.284 = 0.037
D	M	Sí	0.349 x 0.473 x 0.245 = 0.04
D	H	Sí	0.330 x 0.626 x 0.245 = 0.0425
A	M	No	0.176 x 0.115 x 0.289 = 0.005
A	H	No	0.38 x 0.884 x 0.289 = 0.097
B	M	No	0.32 x 0.427 x 0.181 = 0.024
B	H	No	0.37 x 0.957 x 0.181 = 0.064
C	M	No	0.66 x 0.645 x 0.284 = 0.120
C	H	No	0.631 x 0.354 x 0.284 = 0.0634
D	M	No	0.65 x 0.473 x 0.245 = 0.0753
D	H	No	0.67 x 0.526 x 0.245 = 0.086

G	A	P(G, A)
M	Sí	0.302
H	Sí	0.345
M	No	0.224
H	No	0.3104

$$P(A \mid G=M) = \frac{0.302}{(0.302 + 0.224)} = \boxed{0.574}\ \%$$

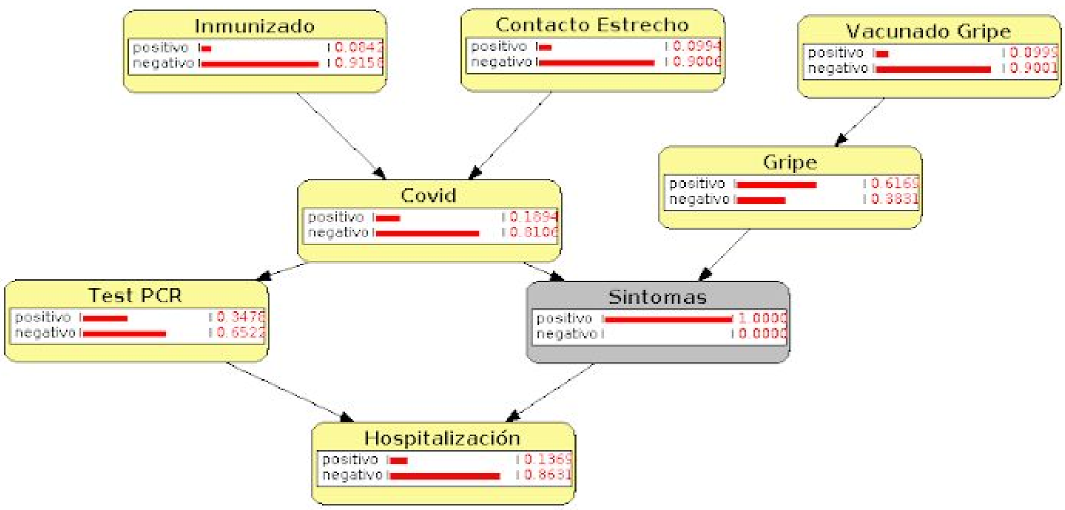
DESARROLLO DE LA PRÁCTICA

1. Red Bayesiana para el Covid

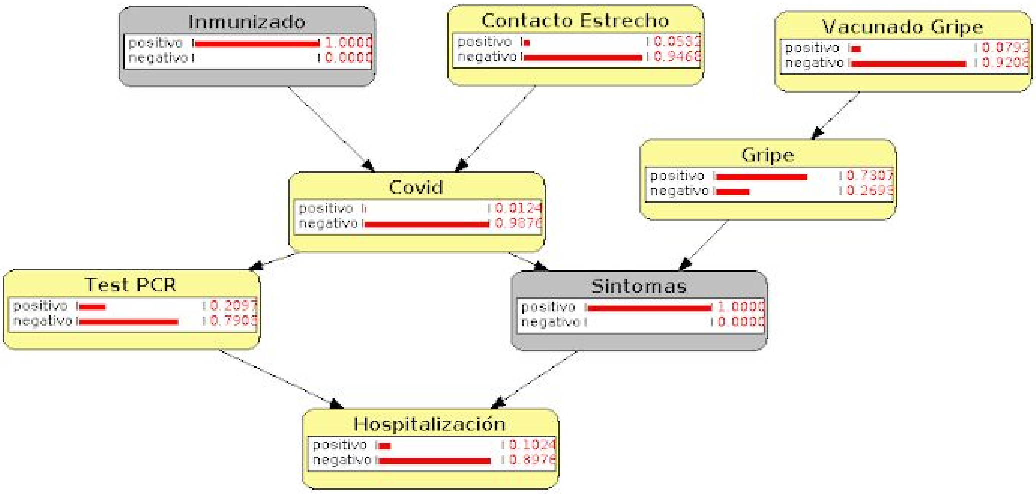


a) Comprobación de independencia condicional.

<u>I y H son independientes dado S</u>
Esta afirmación se ha comprobado en el trabajo previo que es falsa.
Para comprobarlo con OpenMarkov se ha añadido la evidencia de Síntomas positivos



A continuación se ha añadido la evidencia de estar Inmunizado y se ha observado que la tabla de probabilidad de Hospitalización cambia:



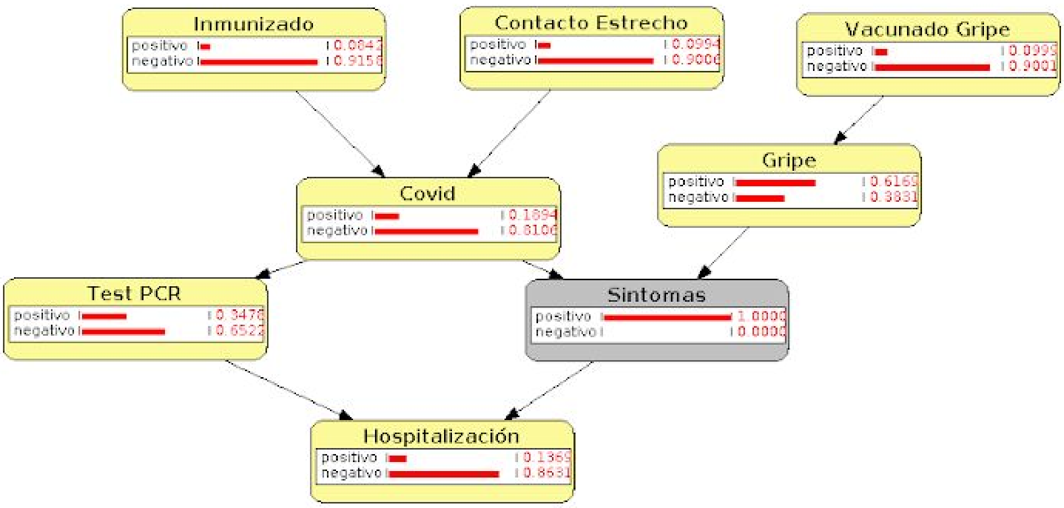
Inicialmente p(+h) = 0.1369, tras añadir la evidencia de estar inmunizado p(+h)=0.1024

Por lo tanto I y H son dependientes dado S.

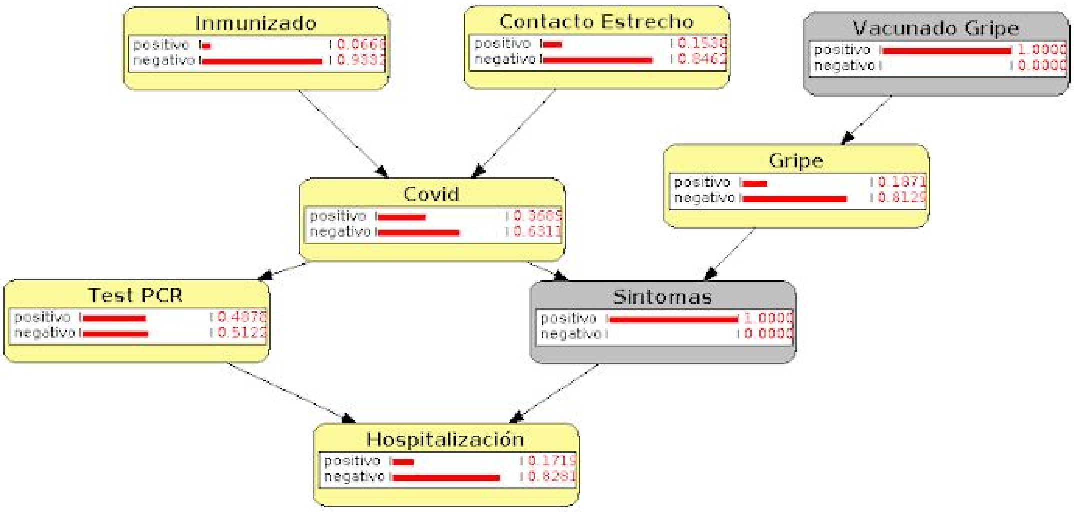
V y H son independientes dado S

Esta afirmación se ha comprobado en el trabajo previo que es falsa.

Para comprobarlo con OpenMarkov se ha añadido la evidencia de Síntomas positivos



A continuación se ha añadido la evidencia de estar vacunado y se ha observado que la tabla de probabilidad de Hospitalización cambia:

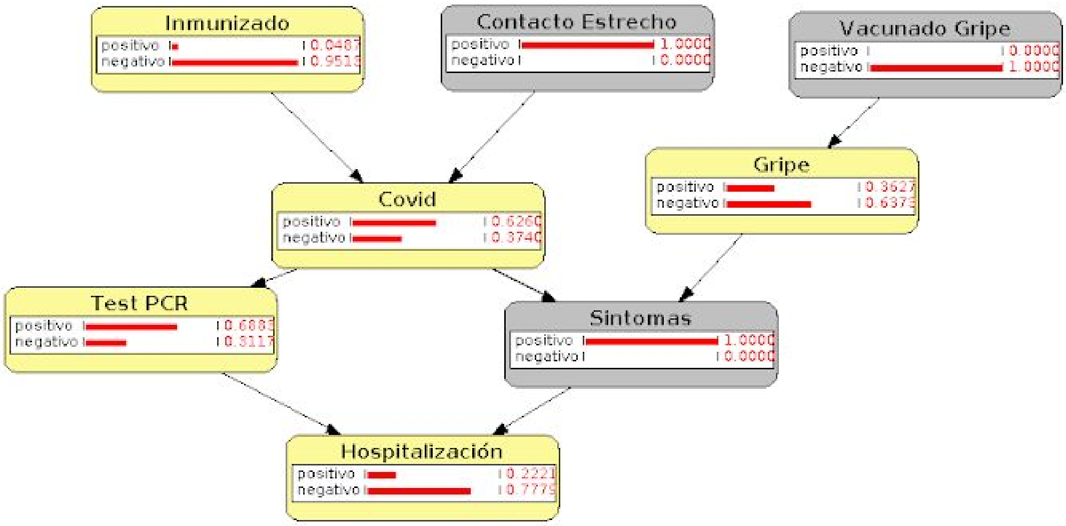


Inicialmente p(+h) = 0.1369, tras añadir la evidencia de estar inmunizado p(+h)=0.1715

Por lo tanto V y H son dependientes dado S.

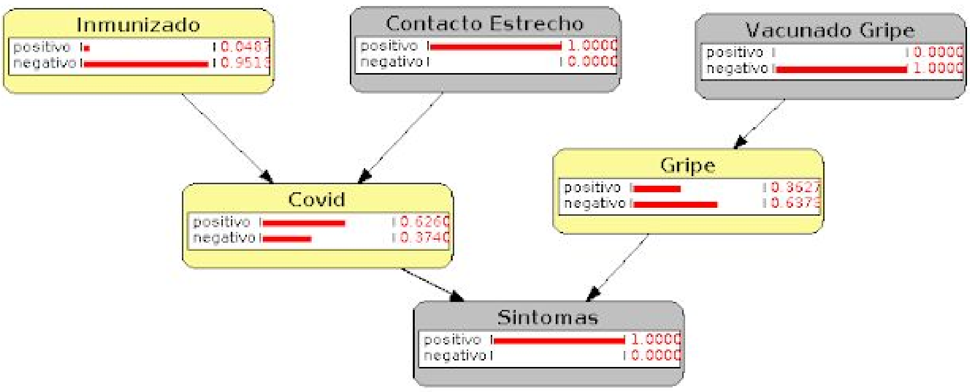
b) Cálculo de la probabilidad p(C | +e, +s, -v)

Se han añadido las evidencias tener un contacto estrecho, tener síntomas y no estar vacunado de la gripe.



Como se puede observar, la probabilidad obtenida es p(C | +e, +s, -v) = 0.6260, tal y como se ha calculado en el trabajo previo.

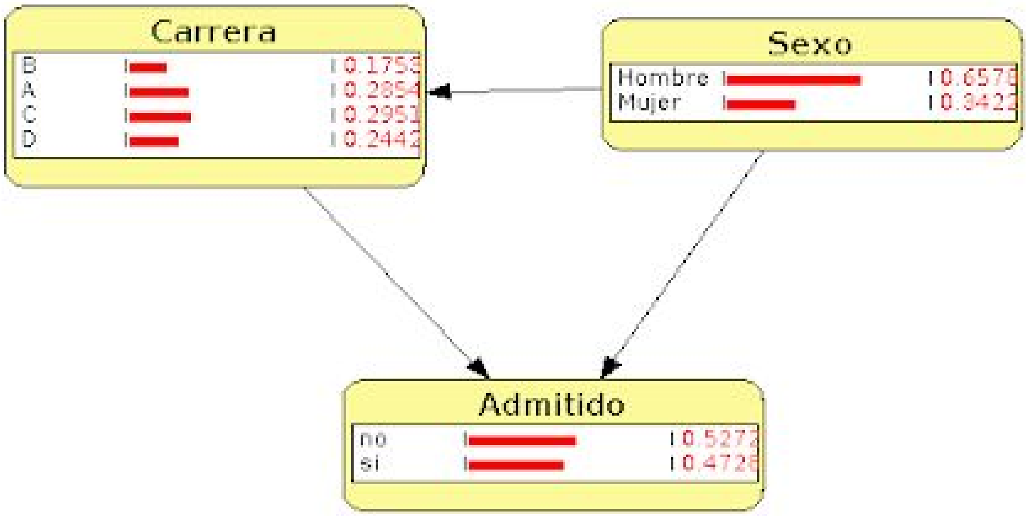
Para comprobar que realmente las variables Test PCR y Hospitalización se pueden ignorar, se han eliminado de la red Bayesiana y se ha observado que la probabilidad p(C | +e, +s, -v) sigue siendo 0.6260.



2. Red Bayesiana Berkley.

a)Comprobación respecto al trabajo previo.

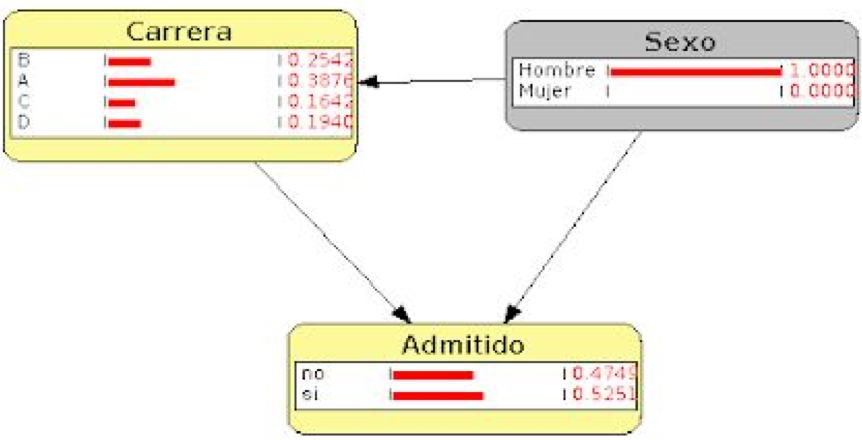
Tras hacer que OpenMarkov aprenda el fichero "DatosBerkeley.xls" se ha obtenido la siguiente red Bayesiana:



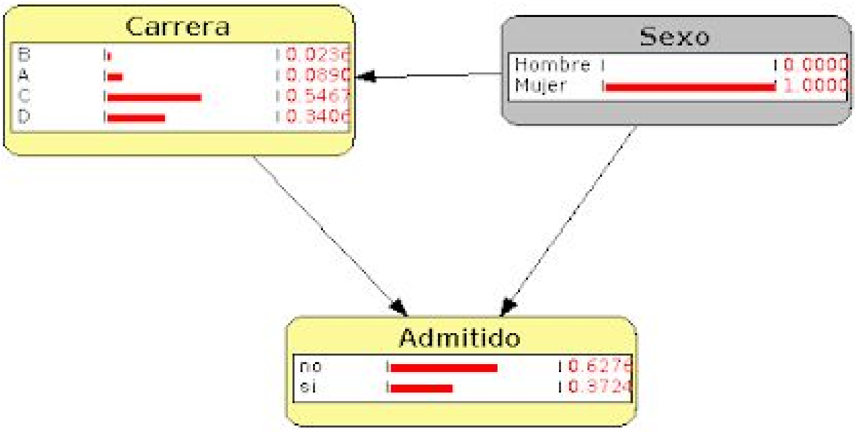
La principal diferencia que observo respecto a mi estudio previo es que la dirección de la causalidad entre Carrera y Sexo la he modelado al revés.
Pese a ello, he observado que en la tabla de probabilidad de Admitido, las probabilidades difieren de mis estudio previo únicamente a partir de las centésimas.

Sexo	Mujer	Mujer	Mujer	Mujer	Hombre	Hombre	Hombre	Hombre
Carrera	D	C	A	B	D	C	A	B
no	0.678938	0.680363	0.186275	0.329268	0.678404	0.620148	0.375784	0.376941
si	0.321062	0.319637	0.813725	0.670732	0.321596	0.379852	0.624216	0.623059

b) Cálculo de probabilidad de admisión para un hombre y para una mujer.
La probabilidad de admisión para un hombre es:



Como se observa en la imagen, p(A|S=Hombre) = 0.5251
La probabilidad de ser admitido siendo una mujer es:



Como se observa en la imagen, p(A|S=Mujer) = 0.3724
(En mi trabajo previo, llego al resultado de que la probabilidad es de 57%. Esta diferencia tan significativa se ha podido deber a modelar “Carrera -> Sexo” en vez de “Carrera <- Sexo” o a algún error en el algoritmo de eliminación.)

c)Si son distintas,¿Hay discriminación de género o puede haber otro motivo?

La probabilidad de admisión siendo mujer es menor del 50%, por lo tanto podría haber discriminación en el proceso de selección.
Sin embargo, cuando añado la evidencia de pertenecer a la carrera A o la carrera B, o carrera C, o carrera D, la probabilidad de ser admitida siento mujer es de 0.81%, 0.67%, 0.31% y 0.32% respectivamente.
Por lo tanto, añadiendo la evidencia de la carrera no queda tan claro que haya discriminación hacia la mujer, sobre todo en las carreras A y B.
Esto podría ser un ejemplo de la paradoja de Simpson presentada en clase.