UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Introducción al entorno de laboratorio y razonamiento probabilístico

DSIC
Departamento de Sistemas Informáticos y Computación

Objetivos formativos

- Introducir el entorno de laboratorio
- Aplicar conceptos y técnicas de razonamiento probabilístico

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Índice

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

1. Introducción al entorno de laboratorio: octave

- Octave es un lenguage interpretado para computo numérico
- Uso interactivo o con ficheros que guardan programas
- Versión libre de MATLAB
- Disponible en `http://www.gnu.org/software/octave`
- Manual de referencia
- Introduciremos octave con ejemplos sobre razonamiento probabilístico
- Inicio de sesión octave: `octave -q`

2. Representación probabilística

El conocimiento probabilístico puede representarse con la distribución de probabilidad conjunta de las variables aleatorias de interés.

Ejemplo del dentista: conocimiento para diagnosticar caries

Variables aleatorias de interés:

$$Dolor: \qquad D \in \{0,1\}$$
$$Caries: \qquad C \in \{0,1\}$$
$$Hueco: \qquad H \in \{0,1\}$$

Representación:

$$P(D = d, C = c, H = h)$$

d	c	h	P
0	0	0	0,576
0	0	1	0,008
0	1	0	0,144
0	1	1	0,072
1	0	0	0,064
1	0	1	0,012
1	1	0	0,016
1	1	1	0,108
	Suma:		1,000

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

La tabla del dentista en octave

d	c	h	P
0	0	0	0,576
0	0	1	0,008
0	1	0	0,144
0	1	1	0,072
1	0	0	0,064
1	0	1	0,012
1	1	0	0,016
1	1	1	0,108

Introduce la tabla del dentista en octave:

```
T = [0 0 0 .576; 0 0 1 .008; 0 1 0 .144; 0 1 1 .072;
     1 0 0 .064; 1 0 1 .012; 1 1 0 .016; 1 1 1 .108];
```

Elemento en la fila 1, columna 4:

```
T(1,4)
```

```
ans = 0.57600
```

Elemento en la fila 1, última columna:

```
T(1,end)
```

```
ans = 0.57600
```

Elementos de la fila 1 a 4 de la última columna:

```
T(1:4,end)
```

```
ans = 0.5760000
      0.0080000
      0.1440000
      0.0720000
```

Elementos (de todas las filas) de la última columna:

```
T(:,end)
```

```
ans = 0.5760000
      0.0080000
      0.1440000
      0.0720000
      0.0640000
      0.0120000
      0.0160000
      0.1080000
```

d	c	h	P
0	0	0	0,576
0	0	1	0,008
0	1	0	0,144
0	1	1	0,072
1	0	0	0,064
1	0	1	0,012
1	1	0	0,016
1	1	1	0,108

Elementos en las filas 1, 2, 5 y 6 de la última columna:

```
T([1 2 5 6],end)
```

```
ans =
    0.5760000
    0.0080000
    0.0640000
    0.0120000
```

Suma de los elementos de la última columna:

```
sum(T(:,end))
```

```
ans =  1.00000
```

Indicadores de filas con elementos nulos en la columna 3:

```
T(:,3)==0
```

```
ans = 1
      0
      1
      0
      1
      0
      1
      0
```

Filas con elementos de la columna 2 no nulos:

```
find(T(:,2))
```

```
ans = 3
      4
      7
      8
```

Filas con elementos nulos en las columnas 2 y 3:

```
find(T(:,2)==0 & T(:,3)==0)
```

```
ans = 1
      5
```

3. Inferencia probabilística

A partir de la distribución conjunta podemos calcular la probabilidad de cualquier *suceso* (*proposición*) mediante aplicación de:

La regla suma:

$$P(x) = \sum_{y} P(x, y)$$

La regla producto:

$$P(x, y) = P(x)\, P(y \mid x)$$

En general no es necesario conocer la tabla completa de probabilidades conjuntas para calcular la probabilidad de un suceso dado.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

d	c	h	P
0	0	0	0,576
0	0	1	0,008
0	1	0	0,144
0	1	1	0,072
1	0	0	0,064
1	0	1	0,012
1	1	0	0,016
1	1	1	0,108

Elementos en última col. de filas con 0 en las cols. 2 y 3:

```
1 T(find(T(:,2)==0 & T(:,3)==0),end)
```

```
1 ans = 0.576000
2       0.064000
```

Probabilidad de caries y hueco (a la vez):

$$P(c=1,h=1)=\sum_{d=0,1}P(d,c=1,h=1)=0{,}180$$

```
1 Pc1h1=sum(T(find(T(:,2)==1 & T(:,3)==1),end))
```

```
1 Pc1h1 = 0.18000
```

Probabilidad de hueco:

$$P(h=1)=\sum_{d=0,1}\sum_{c=0,1}P(d,c,h=1)=0{,}200$$

```
1 Ph1=sum(T(find(T(:,3)==1),end))
```

```
1 Ph1 = 0.20000
```

Probabilidad de caries tras observar (sabiendo que hay) hueco:

$$P(c=1\mid h=1)=\frac{P(c{=}1,h{=}1)}{P(h{=}1)}=\frac{0{,}180}{0{,}200}=0{,}900$$

```
1 Pc1Dh1=Pc1h1/Ph1
```

```
1 Pc1Dh1 = 0.90000
```

Probabilidad de dolor sabiendo que hay caries:

$$P(d=1\mid c=1)=\frac{P(d{=}1,c{=}1)}{P(c{=}1)}=\frac{0{,}124}{0{,}340}=0{,}365$$

```
1 Pd1c1=sum(T(find(T(:,1)==1 & T(:,2)==1),end))
2 Pc1=sum(T(find(T(:,2)==1),end))
3 Pd1Dc1=Pd1c1/Pc1
```

```
1 Pd1c1  = 0.12400
2 Pc1    = 0.34000
3 Pd1Dc1 = 0.36471
```

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

4. Ejercicio: aplicación del teorema de Bayes

El ***teorema de Bayes*** permite actualizar nuestro conocimiento sobre una hipótesis y tras observar una nueva evidencia x:

$$P(y \mid x) = \frac{P(x,y)}{P(x)} = P(y)\,\frac{P(x \mid y)}{P(x)}$$

De otra forma: $P(y \mid x)$ es la probabilidad de que se produzca el efecto y tras observar que se ha producido la causa x.

Ejercicio: calcula la probabilidad de caries sabiendo que hay dolor

$$P(c=1 \mid d=1) = P(c=1)\,\frac{P(d=1 \mid c=1)}{P(d=1)}$$

Pd1 = 0.20000
\>> Pd1c1
Pd1c1 = 0.12400
\>> Pc1Dd1 = Pd1c1 / Pd1
Pc1Dd1 = 0.62000

UNIVERSITAT POLITÈCNICA DE VALÈNCIA