

Curso de Experto Universitario en
Probabilidad y Estadística en Medicina

www.ia.uned.es/cursos/prob-estad

Método bayesiano clásico

F. J. Díez Vegas

Dpto. Inteligencia Artificial. UNED

fjdiez@dia.uned.es

www.ia.uned.es/~fjdiez

Método bayesiano clásico

◆ Diagnósticos

- d_1 : Insuficiencia mitral crónica
- d_2 : Insuficiencia mitral aguda
- d_3 : No hay insuficiencia mitral

◆ Posibles hallazgos

➤ H_1 : Disnea

- $h_{1,1}$: No hay disnea
- $h_{1,2}$: Disnea leve
- $h_{1,3}$: Disnea moderada
- $h_{1,4}$: Disnea severa

➤ H_2 : Dilatación de la A.I.

- $h_{2,1}$: No dilatada
- $h_{2,2}$: Dilatación leve
- $h_{2,3}$: Dilatación moderada
- $h_{2,4}$: Dilatación severa

➤ H_3 : Regurgitación observada en eco Doppler

- $h_{3,1}$: ...

Fundamentos del método bayesiano clásico

- ◆ Punto de partida:
 n diagnósticos, m hallazgos

- ◆ Parámetros:
 - > 2^n probabilidades a priori:
 $P(d_1, \dots, d_n)$
 - > 2^{n+m} probabilidades condicionadas:
 $P(h_1, \dots, h_m | d_1, \dots, d_n)$

- ◆ Teorema de Bayes

$$P(d_1, \dots, d_n | h_1, \dots, h_m) = \frac{P(h_1, \dots, h_m | d_1, \dots, d_n) \cdot P(d_1, \dots, d_n)}{\sum_{d'_1, \dots, d'_n} P(h_1, \dots, h_m | d'_1, \dots, d'_n) \cdot P(d'_1, \dots, d'_n)}$$

- ◆ Número de parámetros:
 - > 3 diagnósticos, 10 hallazgos → 8.191 parámetros
 - > 10 diagnósticos, 50 hallazgos → $1_3 152.921_2 504.606_1 846.975$ parám.

1ª hipótesis: diagnósticos exclusivos y exhaustivos

- ◆ Diagnósticos exclusivos: el paciente sólo tiene una enfermedad
Diagnósticos exhaustivos: cubren todas las posibilidades
- ◆ Resultado: variable D que representa los diagnósticos posibles
- ◆ Parámetros:

- > n probabilidades a priori:
 $P(d_j)$
- > $2^{m \cdot n}$ probabilidades condicionadas:
 $P(h_1, \dots, h_m | d_j)$

- ◆ El teorema de Bayes se reduce a

$$P(d_j | h_1, \dots, h_m) = \frac{P(h_1, \dots, h_m | d_j) \cdot P(d_j)}{\sum_j P(h_1, \dots, h_m | d_j) \cdot P(d_j)}$$

- ◆ Número de parámetros:
 - > 3 diagnósticos, 10 hallazgos → 4.095 parámetros
 - > 10 diagnósticos, 50 hallazgos → $12.384_2 898.975_1 268.863$ parám.

2ª hipótesis: independencia condicional

- ◆ Supone que, para cada diagnóstico, los hallazgos son condicionalmente independientes entre sí:

$$P(h_1, \dots, h_m | d_i) = P(h_1 | d_i) \cdot \dots \cdot P(h_m | d_i)$$

- ◆ Parámetros:

- > n probabilidades a priori:

$$P(d_i)$$

- > $m \cdot n$ probabilidades condicionadas:

$$P(h_j | d_i)$$

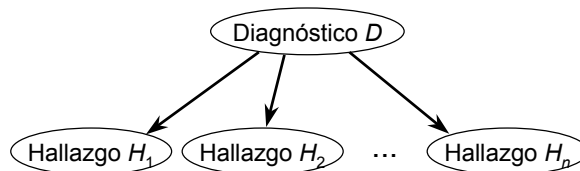
- ◆ Teorema de Bayes

$$P(d_i | h_1, \dots, h_m) = \frac{P(h_1 | d_i) \cdot \dots \cdot P(h_m | d_i) \cdot P(d_i)}{\sum_j P(h_1 | d_j) \cdot \dots \cdot P(h_m | d_j) \cdot P(d_j)}$$

- ◆ Número de parámetros:

- > 3 diagnósticos, 10 hallazgos → 43 parámetros
- > 10 diagnósticos, 50 hallazgos → 560 parámetros

- ◆ Representación gráfica



- ◆ Comparación de dos diagnósticos

$$\frac{P(d_j | h_1, \dots, h_m)}{P(d_i | h_1, \dots, h_m)} = \frac{P(d_j)}{P(d_i)} \cdot \frac{P(h_1 | d_j)}{P(h_1 | d_i)} \cdot \dots \cdot \frac{P(h_m | d_j)}{P(h_m | d_i)}$$

- > Comparar $P(\text{enfermedad-1})$ frente a $P(\text{enfermedad-2})$
- > Comparar $P(+e)$ frente a $P(-e)$:

$$RP_{post} = RP_{pre} \cdot RV_1 \cdot \dots \cdot RV_m$$

Ejemplo: Diagnóstico de una patología

◆ Enfermedad E

➤ Prevalencia: $P(+e)=0'002$

◆ Síntoma S

➤ Sens: $P(+s|+e)=0'93$

Espec: $P(-s|-e)=0'99$

➤ $RV_{+s} : 0'93/(1-0'99)=93'00$

$RV_{-s} : (1-0'93)/0'99=0'0707$

◆ Prueba analítica A

➤ Sens: $P(+a|+e)=0'995$

Espec: $P(-a|-e)=0'997$

➤ $RV_{+a} : 0'995/(1-0'997)=331'6$

$RV_{-a} : (1-0'995)/0'997=0'005$

◆ Diagnóstico

$$P(+e|s,a) = \frac{P(s|+e) \cdot P(a|+e) \cdot P(+e)}{P(s|+e) \cdot P(a|+e) \cdot P(+e) + P(s|-e) \cdot P(a|-e) \cdot P(-e)}$$

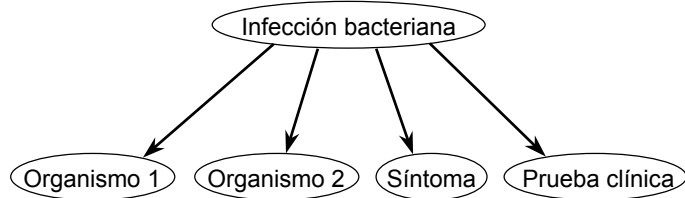
$$RP_{post} = RP_{pre} \cdot RV_S \cdot RV_A$$

Ejemplo: Resultados

| Evid | RP_{pre} | RV_S | RV_A | RP_{post} | $P(+e evid)$ |
|--------|------------|--------|--------|-----------------------|-----------------------|
| — | 0'002 | 1 | 1 | 0'0002 | 0'002 |
| +s | 0'002 | 93'000 | 1 | 0'1860 | 0'157 |
| -s | 0'002 | 0'0707 | 1 | 1'41·10 ⁻⁴ | 1'41·10 ⁻⁴ |
| +a | 0'002 | 1 | 331'6 | 0'6632 | 0'399 |
| -a | 0'002 | 1 | 0'005 | 1'00·10 ⁻⁵ | 1'00·10 ⁻⁵ |
| +s, +a | 0'002 | 93'000 | 331'6 | 61'678 | 0'984 |
| +s, -a | 0'002 | 93'000 | 0'005 | 0'00093 | 0'00093 |
| -s, +a | 0'002 | 0'0707 | 331'6 | 0'047 | 0'045 |
| -s, -a | 0'002 | 0'0707 | 0'005 | 7'10·10 ⁻⁷ | 7'10·10 ⁻⁷ |

Inconvenientes del método bayesiano clásico

- ◆ Hipótesis de diagnósticos exclusivos:
Supone que el paciente sólo tiene una enfermedad
- ◆ Hipótesis de independencia condicional



Cuándo se puede aplicar el método bayesiano clásico

- ◆ Diagnósticos exclusivos:
 - Diagnosticar una sola enfermedad (presente / ausente)
 - Varios diagnósticos, pero es muy improbable que una persona tenga dos enfermedades simultáneamente
- ◆ Independencia condicional
 - Varios efectos de una enfermedad, con mecanismos causales independientes

