Fórmula de Bayes: caso de estudio "tigre oculto", número arbitrario de rugidos, NO recursivo

© 2024, Antonio Sala Piqueras. Universitat Politecnica de Valencia. Todos los derechos reservados.

Presentaciones en vídeo: http://personales.upv.es/asala/YT/V/tiger4.html http://personales.upv.es/asala/YT/V/tiger5.html

Este código ejecutó en Matlab **R2023a**

Objetivos: Ilustrar la fórmula de Bayes en el caso "*Tigre oculto detrás de dos puertas*", donde se deben acumular mediciones ruidosas sobre la ubicación del tigre para mejorar las posibilidades de supervivencia (evaluar correctamente dónde está el tigre antes de "actuar" abriendo una puerta).

El "decidir" la acción óptima NO será objetivo de este material, centrado sólo en actualizar la "creencia" (belief en inglés) conforme se reciben observaciones.

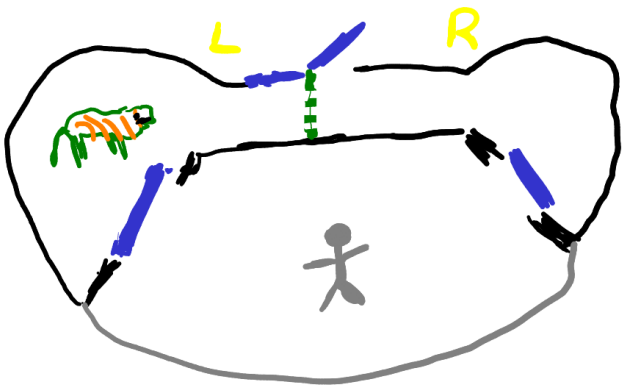
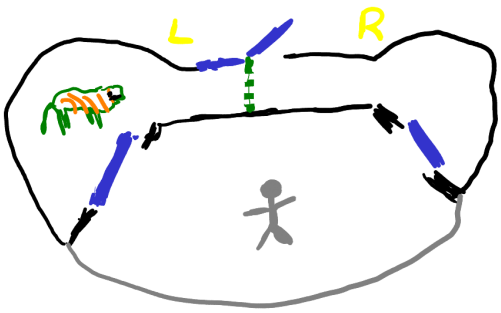


Tabla de Contenidos

Planteamiento del problema 1
Construcción/Interpretación de tabla de probabilidad 2
Caso 0: Se cierra reja del pasillo central y escucho UN rugido 3
Actualización de la creencia (posterior): fórmula de Bayes 4
Revisión fórmula de Bayes, notación compacta 5
Fórmula de Bayes para DOS rugidos (NO recursiva) 6
Fórmula de Bayes para número arbitrario de rugidos (NO recursiva) 9

Planteamiento del problema

Tenemos una situación donde un empleado de zoológico debe "adivinar" dónde está el tigre, para escapar por el lado donde no esté sin ser devorado por él:



El tigre puede estar a la izquierda o a la derecha, y sus rugidos pueden escucharse "más fuerte" por la puerta izquierda o por la derecha. Simplificamos a:

Medidas (observaciones) binarias: $m \in \{$ `hear_left` (1, HL), `hear_right` (2, HR) $\}$.

Estados tigre binarios: $t \in \{$ `tiger_left` (1, TL), `tiger_right` (2, TR) $\}$.

```
hear_L=1;  hear_R=2; %El número únicamente significa posición en listas
tiger_L=1; tiger_R=2; %El número únicamente significa posición en listas
```

Construcción/Interpretación de tabla de probabilidad

Definimos matriz PM, donde PM(m,e) es la probabilidad de medir "m" cuando el estado del tigre es "t". O sea, es una tabla de probabilidad **condicional**.

$$\text{PM= [} p(HL|TL) \quad p(HL|TR) \text{ ;}$$

$$p(HR|TL) \quad p(HR|TR)]$$

Filas representan "observaciones", columnas representan "estados".

Propiedades: COLUMNAS suman 1. No tiene por qué ser simétrica.

```
%PM=[.99 .9;.01 .1]; %non symmetric... different measurement noise at each state... mos
%PM=[.8 .1;.2 .9];
%PM=[.8 .2;.2 .8];
PM=[.6 .4;.4 .6];
%PM=[.5 .0001;.5 .9999];
PM
```

```
PM = 2×2
    0.6000    0.4000
    0.4000    0.6000
```

Será nuestro "modelo del mundo tigre" y haremos "inferencia estadística" a partir de él (estimar dónde está el tigre, decidir qué puerta abrir, decidir cuántos rugidos escuchar hasta estar "suficientemente seguros" para abrir la puerta -- eso depende del miedo que tenga, que se modela con funciones de utilidad/coste adicionales--)...

Caso 0: Se cierra reja del pasillo central y escucho UN rugido

La creencia ("belief") se denotará con b. Será la probabilidad "a priori" de `tiger_right`. Obviamente, la probabilidad a priori de `tiger_left` será $1-b$.

Si lo hacemos "en letra" lo de la marginal, tenemos:

```
syms b real % b := p(TR) a priori
prob_hearL(b)= PM(hear_L,tiger_R)*b + PM(hear_L,tiger_L)*(1-b);
vpa(prob_hearL)
```

ans(b) = $0.6 - 0.2\,b$

```
prob_hearR(b)= PM(hear_R,tiger_R)*b + PM(hear_R,tiger_L)*(1-b);
vpa(prob_hearR)
```

ans(b) = $0.2\,b + 0.4$

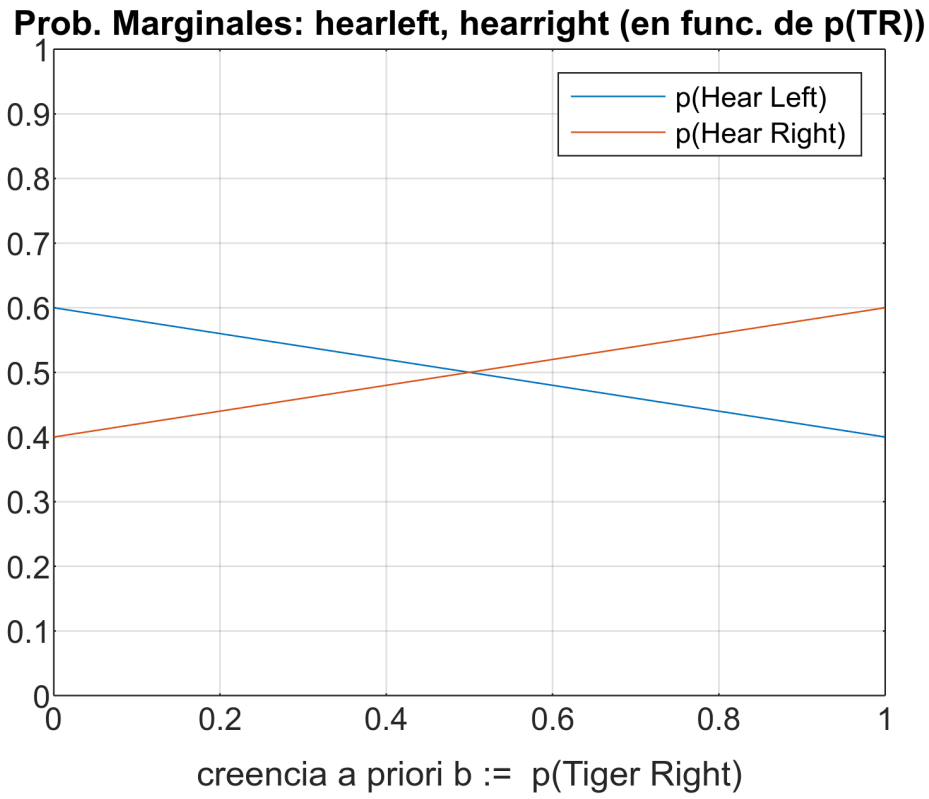
En notacion matricial resulta "matriz de probabilidades condicionales" x "vector de probabilidades a priori":

```
vpa(  PM*[(1-b); b]  , 4)
```

ans =

$$\begin{pmatrix} 0.6 - 0.2\,b \\ 0.2\,b + 0.4 \end{pmatrix}$$

```
fplot([prob_hearL;prob_hearR],[0 1]), grid on, ylim([0 1])
legend("p(Hear Left)","p(Hear Right)") %OJO interpretacion, b no es "evento"...
xlabel("creencia a priori b :=  p(Tiger Right)"), title("Prob. Marginales: hearleft, he
```



Actualización de la creencia (posterior): fórmula de Bayes

Si midiese `hear_left`, lo podemos escribir a partir de la tabla "condicional" y la probabilidad "a priori" del lado donde se situa el tigre:

$$p(TR|hearleft) = \frac{p(hearleft|TR) \cdot p(TR)}{p(hearleft|TR) \cdot p(TR) + p(hearleft|TL) \cdot p(TL)} = \frac{p(hearleft|TR) \cdot}{p(hearleft|TR) \cdot b + p(hearle}$$

Antes lo hemos hecho para un valor de b numérico dado... ahora lo escribo en simbólico:

```
posteriori_ifHL(b) = PM(hear_L,tiger_R)*b/prob_hearL(b); %dividimos por la marginal
vpa(posteriori_ifHL)
```

```
ans(b) =
```

$$-\frac{0.4\,b}{0.2\,b - 0.6}$$

Si midiese `hear_right`,

$$p(TR|hearright) = \frac{p(hearright|TR) \cdot p(TR)}{p(hearright|TR) \cdot p(TR) + p(hearright|TL) \cdot p(TL)} = \frac{p(hearrig}{p(hearright|TR) \cdot b + p}$$

```
posteriori_ifHR(b) = PM(hear_R,tiger_R)*b/prob_hearR(b);%dividimos por la marginal
```

```
vpa(posteriori_ifHR)
```

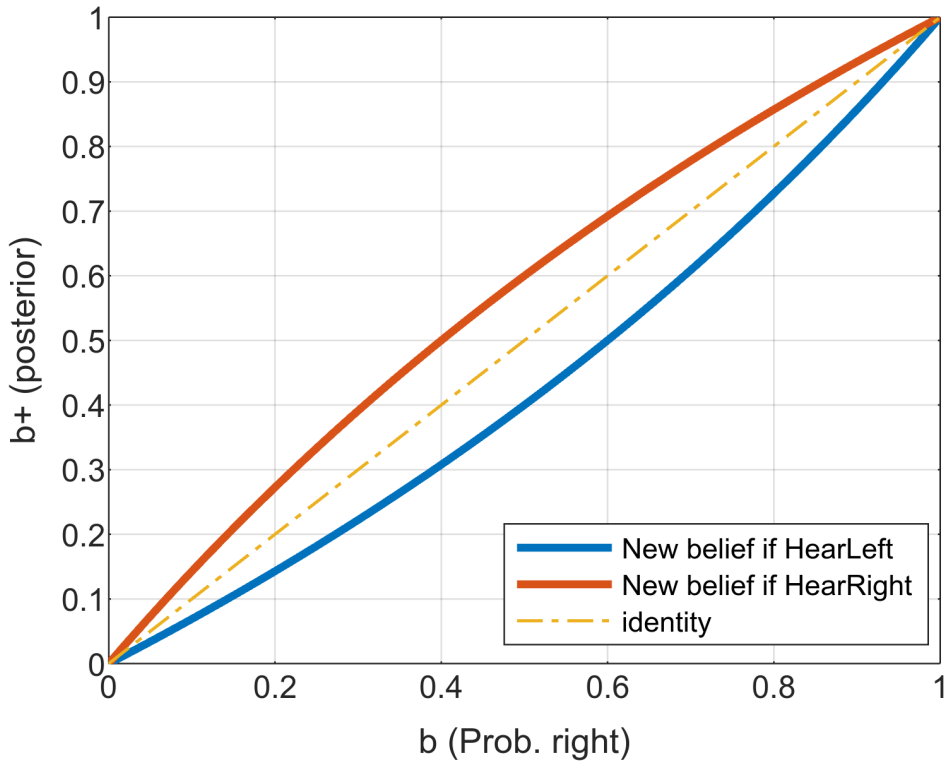
```
ans(b) =
```

$$\frac{0.6\,b}{0.2\,b+0.4}$$

Representación gráfica:

El proceso "conceptual" es que a partir de una "creencia" (probabilidad a priori) y una observación, calculo una "creencia actualizada" (probabilidad a posteriori): $b_+ = Bayes(b,m)$. $b \in [0,1]$, $m \in \{HL, HR\}$. Gráficamente:

```
fplot([posteriori_ifHL;posteriori_ifHR],[0 1],LineWidth=2.5), grid on
hold on, fplot(b,[0 1],'-.',LineWidth=1), hold off
xlabel('b (Prob. right)'), ylabel('b+ (posterior)')
legend("New belief if HearLeft","New belief if HearRight",'identity',Location="best")
```



No hay un punto "fijo" en el que la creencia nueva sea igual a la antigua excepto la creencia total a la derecha (b=1) o izquierda (b=0). Si estoy "100% seguro" de algo, experimentar no va a aportar nueva información.

Revisión fórmula de Bayes, notación compacta

Llamemos, por compactar notación

$$\text{PM}= \begin{pmatrix} p(HL|TL) & p(HL|TR) \\ p(HR|TL) & p(HR|TR) \end{pmatrix} = \begin{pmatrix} q & r \\ 1-q & 1-r \end{pmatrix}$$

a las condicionales HL HR (filas) dado TL y TR (columnas), respectivamente. Si $b = p(TR)$, la fórmula de Bayes dirá que, según observe HL o HR, aplicaré una de estas dos fórmulas:

$$b_+ = p(TR \mid HL) = \frac{r \cdot b}{q \cdot (1-b) + r \cdot b}$$

$$b_+ = p(TR \mid HR) = \frac{(1-r) \cdot b}{(1-q) \cdot (1-b) + (1-r) \cdot b}$$

Fórmula de Bayes para DOS rugidos (NO recursiva)

Si me espero hasta escuchar DOS rugidos (en abstracto observación $o_1 \in \{HL, HR\}$ y $o_2 \in \{HL, HR\}$), suponiendo "independencia estadística" (bueno, en rigor se denomina "independencia condicional" porque estamos hablando de productos de probabilidades condicionales), entonces las nuevas condicionales respecto a TL son:

$$p((HL, HL)|TL) \text{ es } \quad p(HL|TL) \cdot p(HL|TL) = q^2,$$

$$p((HL, HR)|TL) = p((HR, HL)|TL) = p(HL|TL)p(HR|TL) = q \cdot (1-q) \ ,$$

$$p((HR, HR)|TL) = p(HR|TL)p(HR|TL) = (1-q)^2$$

y lo mismo con las condicionales respecto a TR:

$$p((HL, HL)|TR) \text{ es } r^2,$$

$$p((HL, HR)|TR) = P((HR, HL)|TR) = r \cdot (1-r) \ ,$$

$$p((HR, HR)|TR) = (1-r)^2$$

Con lo cual mi nueva tabla condicional para DOS rugidos para aplicar Bayes será:

$$\text{PM2} = \begin{pmatrix} p(HLL|TL) & p(HLL|TR) \\ p(HLR|TL) & p(HLR|TR) \\ p(HRL|TL) & p(HRL|TR) \\ p(HRR|TL) & p(HRR|TR) \end{pmatrix} = \begin{pmatrix} q^2 & r^2 \\ q(1-q) & r(1-r) \\ q(1-q) & r(1-r) \\ (1-q)^2 & (1-r)^2 \end{pmatrix}$$

*HLL es una abreviatura de ($o_1 =$ HL, $o_2 =$HL) y similar para las otras combinaciones de observaciones.

Nota: El significado de $o_2 \wedge o_1$ se refiere a observaciones en distinto instante de tiempo... No podemos observar $o_1 = HL$, $o_2 = HR$ si se refiriesen al mismo instante, obviamente. Los problemas que estudian variables (aleatorias) en diferentes instantes de tiempo se dice que estudian la "dinámica" de un sistema.

Por ello, formalmente, debería ser $o_1 \in \{HL_{t=1}, HR_{t=1}\}$ y $o_2 \in \{HL_{t=2}, HR_{t=2}\}$, aunque omitiremos el subíndice temporal asociado a cada evento por simplicidad de notación.

Fórmula de Bayes para la observación $(HL_{t=1}, HL_{t=2})$... que abrevio a HLL:

$$b_{++} = p(TR \mid HLL) = \frac{p(HLL \mid TR)p(TR)}{p(HLL \mid TL)p(TL) + p(HLL \mid TR)p(TR)}$$

$$= \frac{r^2 \cdot b}{q^2 \cdot (1-b) + r^2 \cdot b}$$

deberíamos completar todas las opciones:

$$p(TR \mid HLL), p(TR \mid HLR), p(TR \mid HRL), p(TR \mid HRR)$$

*Se deja como ejercicio.

las $p(TL|\dots)$ serán 1-la correspondiente de $p(TR|\dots)$...Esas no hacen falta guardar.

Para 3 rugidos (sin que el tigre pueda cambiar de lado, reja central cerrada), suponiendo "independencia condicional", la tabla para la fórmula de Bayes cuando observemos algo sería:

$$\text{PM3}=\begin{pmatrix} p(HLLL|TL) & p(HLLL|TR) \\ p(HLLR|TL) & p(HLLR|TR) \\ p(HLRL|TL) & p(HLRL|TR) \\ p(HLRR|TL) & p(HLRR|TR) \\ p(HRLL|TL) & p(HRLL|TR) \\ p(HRLR|TL) & p(HRLR|TR) \\ p(HRRL|TL) & p(HRRL|TR) \\ p(HRRR|TL) & p(HRRR|TR) \end{pmatrix} = \begin{pmatrix} q^3 & r^3 \\ q^2(1-q) & r^2(1-r) \\ q^2(1-q) & r^2(1-r) \\ q(1-q)^2 & r(1-r)^2 \\ q^2(1-q) & r^2(1-r) \\ q(1-q)^2 & r(1-r)^2 \\ q(1-q)^2 & r(1-r)^2 \\ (1-q)^3 & (1-r)^3 \end{pmatrix}$$

Reducción de tabla "sólo contar"

Si los eventos son "contar 2L y 1R" (el orden no importa), entonces podemos resumir la tabla en:

$$\text{PM3b}=\begin{pmatrix} p(HL^3|TL) & p(HL^3|TR) \\ p(HL^2R|TL) & p(HL^2|TR) \\ p(HLR^2|TL) & p(HLR^2|TR) \\ p(HR^3|TL) & p(HR^3|TR) \end{pmatrix} = \begin{pmatrix} q^3 & r^3 \\ 3q^2(1-q) & 3r^2(1-r) \\ 3q(1-q)^2 & 3r(1-r)^2 \\ (1-q)^3 & (1-r)^3 \end{pmatrix}$$

Las fórmulas de Bayes acabarán dando el mismo resultado: dividir cada elemento de la fila asociada a la observación experimentada por la suma de la fila resulta lo mismo esté o no multiplicada por 3. Ello hace que, en este tipo de problemas, el orden de las observaciones sea obviado en la mayoría de casos, porque así se reduce el número de posibles eventos y las tablas son más pequeñas (¡imagina lo grande que sería todo con 20 rugidos!).

Fórmula de Bayes para número arbitrario de rugidos (NO recursiva)

Con la combinatoria adecuada, todo se puede resumir en que si de un total de $m+n$ rugidos escuchamos m por la izquierda y n por la derecha, haríamos Bayes con:

$$\begin{pmatrix} p(HL^mR^n \mid TL) & p(HL^mR^n \mid TR) \end{pmatrix} = \begin{pmatrix} \frac{(m+n)!}{m!\,n!}q^m(1-q)^n & \frac{(m+n)!}{m!\,n!}r^m(1-r)^n \end{pmatrix}$$

Los libros de estadística (busca distribución "***binomial***" en tus libros de texto) suelen hacerlo así.

Dirán que las probabilidades condicionales tienen distribución binomial, en concreto:

- el número de rugidos "izquierda" de un total de "Tot" observaciones (condicionadas a **TL**) es una variable aleatoria con distribución binomial de parámetros (Tot, q).
- el número de rugidos "izquierda" de un total de "Tot" observaciones (condicionadas a **TR**) es una variable aleatoria con distribución binomial de parámetros (Tot, r).

Ejemplo numérico:

Leamos la tabla de probabilidad condicional con la que comenzamos todo:

```
PM
```

```
PM = 2×2
    0.6000    0.4000
    0.4000    0.6000
```

```
q=PM(hear_L,tiger_L)
```

```
q = 0.6000
```

```
r=PM(hear_L,tiger_R)
```

```
r = 0.4000
```

Si escucho rugidos "m" veces a la izquierda y "n" a la derecha, entonces:

```
m=4; %number of HL events, num. rugidos izquierda
n=4; %number of HR events, num. rugidos derecha
Tot=m+n; %total roaring events, total rugidos
p_LmRn_cond_TL = nchoosek(Tot,m) * q^m * (1-q)^n
```

```
p_LmRn_cond_TL = 0.2322
```

```
p_LmRn_cond_TR = nchoosek(Tot,m) * r^m * (1-r)^n
```

```
p_LmRn_cond_TR = 0.2322
```

Con lo que aplicando Bayes, la probabilidad a posteriori de que el tigre esté a la izquierda o a la derecha será:

```
b=0.5; %probabilidad a priori de TR
p_TL_cond_LmRn = p_LmRn_cond_TL*(1-b) / ( p_LmRn_cond_TL*(1-b) + p_LmRn_cond_TR*b )
```

```
p_TL_cond_LmRn = 0.5000
```

```
p_TR_cond_LmRn = p_LmRn_cond_TR*b / ( p_LmRn_cond_TL*(1-b) + p_LmRn_cond_TR*b)
```

```
p_TR_cond_LmRn = 0.5000
```