Probabilidades marginal/conjunta/condicional/priori: caso de estudio "tigre oculto"

Presentaciones en vídeo:

http://personales.upv.es/asala/YT/V/tiger1.html

http://personales.upv.es/asala/YT/V/tiger2.html

Este código ejecutó en Matlab **R2023a**

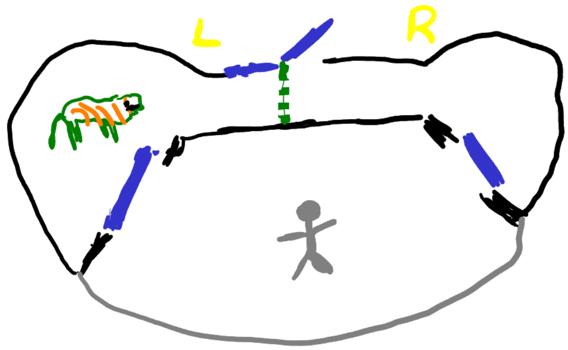
Objetivos: Ilustrar ideas preliminares en el caso "*Tigre oculto detrás de dos puertas*", donde se deben acumular mediciones ruidosas sobre la ubicación del tigre para mejorar las posibilidades de supervivencia (evaluar correctamente dónde está el tigre antes de "actuar" abriendo una puerta).

Este material se centra en las "tablas de probabilidad" y sus diferentes tipos.

Tabla de Contenidos

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,t) es la probabilidad de medir "m" cuando el estado del tigre es "t". O sea, es una tabla de probabilidad **condicional**.

PM= [$p(HL|TL)$ $p(HL|TR)$;

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

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... most times I hear "left" even if it is "right".
PM=[.8 .1;.2 .9];
%PM=[.65 .35;.35 .65];
PM=[.5 .0001;.5 .9999];
PM
```

```
PM = 2×2
    0.5000    0.0001
    0.5000    0.9999
```

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--)...

Probabilidad conjunta versus condicional

Teniamos una tabla de probabilidad **condicional**.

PM= [$p(HL|TL)$ $p(HL|TR)$;

$p(HR|TL)$ $p(HR|TR)$] ;

La tabla de probabilidad "CONJUNTA" es la tabla de probabilidad de los 4 eventos "individuales" posibles... Considerados como "una variable aleatoria que puede tomar 4 valores":

[(HL$\wedge$TL), (HL$\wedge$TR), (HR$\wedge$TL), (HR$\wedge$TR)]

Considerados como una "variable aleatoria bidimensional", los podemos poner en tabla 2x2, con probabilidades que denotaremos:

PCONJUNTA= [*p*(HL,TL) *p*(HL, TR) ;

p(HR,TL) *p*(HR,TR)];

Hemos cambiado la conectiva "y" ($\wedge$) por valores 2D (a,b)... pero es en esencia lo mismo.

Pero por el momento NO PODEMOS escribir dicha tabla conjunta: FALTA UN DATO.

La prob. condicional se define como $p(m|t) := \dfrac{p(m \wedge t)}{p(t)}$, o sea: $p(m \wedge t) = p(m|t) \cdot p(t)$.

Necesitamos saber el valor $p(TL)$, $p(TR)$ ... obviamente suman uno, de modo que con saber $b := p(TR)$ tendremos $p(TL) = 1 - b$.

En efecto, entonces:

PCONJUNTA= [$p(HL|TL) \cdot (1-b) \quad p(HL|TR) \cdot b$;

$p(HR|TL) \cdot (1-b) \quad p(HR|TR) \cdot b$] ;

```
b=0.9; %b=0.5 would be a non-informative prior
PCONJUNTA=[PM(:,1)*(1-b) PM(:,2)*b]
```

```
PCONJUNTA = 2×2
    0.0500    0.0001
    0.0500    0.8999
```

```
sum(sum(PCONJUNTA))
```

```
ans = 1
```

```
PM %condicional, por tenerla a la vista
```

```
PM = 2×2
    0.5000    0.0001
    0.5000    0.9999
```

Interpretacion de $pconjunta(m,e)$**:** Es la probabilidad de una cierta medida (observación) y de un cierto lugar donde el tigre se halle... considerando a las dos como variables aleatorias... "tirando el dado a la vez"...

Pero, ¿se puede interpretar esa probabilidad como "porcentaje de veces que ocurre el evento si se repite muchas veces"?

- **Sí y no:**

-- si repetir significa que "se abre la puerta de paso intermedia y se deja pasar mucho tiempo (tigre cambia de lado si quiere), y se escucha UNA única vez el rugido" entonces **SÍ**;

-- si repetir significa "se cierra la puerta de paso intermedia para que el tigre esté siempre a un lado y se espera hasta escuchar 10 rugidos" entonces **NO**. En esa situación (o si abro puerta pero dejo pasar "poco tiempo" antes de cerrar otra vez), la probabilidad de que el tigre esté en cierto lado depende de dónde estaba "un minuto antes"... y los "rugidos pasados" me dan información sobre dónde podría estar en el pasado que se podria extrapolar al instante actual. Entonces $pconjunta(m,t)$ $p(m|t)$ y $p(t)$ requieren de una interpretación/utilización diferente.

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 $1 - b$.

Probabilidad de las observaciones (marginal)

```
PCONJUNTA
```

```
PCONJUNTA = 2×2
    0.0500    0.0001
    0.0500    0.8999
```

```
marg_TLTR=sum(PCONJUNTA) %suma por columnas, marginal de "tigre" (suma de
elementos de cada columna)
```

```
marg_TLTR = 1×2
    0.1000    0.9000
```

```
marg_HLHR=sum(PCONJUNTA,2) %suma por filas [la dimensión 2], marginal de
"escuchar"
```

```
marg_HLHR = 2×1
    0.0501
    0.9499
```

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) = $0.5 - 0.4999\,b$

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

ans(b) = $0.4999\,b + 0.5$

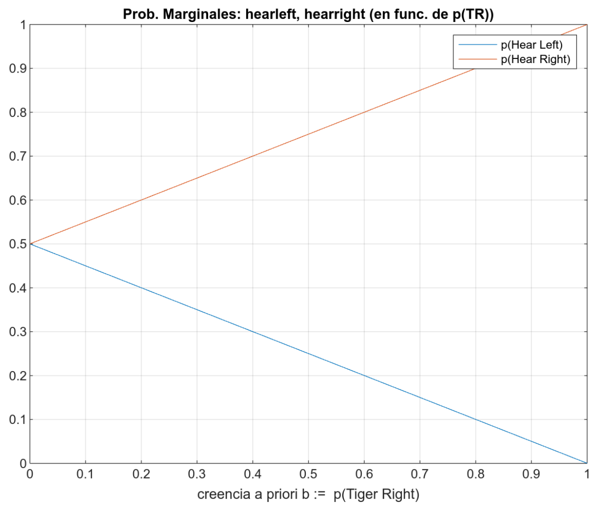
En notacion matricial resulta "matriz de probabilidades condicionales" x "vector de probabilidades a priori":

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

ans =

$$\begin{pmatrix} 0.5 - 0.4999\,b \\ 0.4999\,b + 0.5 \end{pmatrix}$$

```
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, hearright (en func. de p(TR))")
```



Podemos recuperar las "condicionales" de la conjunta:

```
PCONJUNTA
```

```
PCONJUNTA = 2×2
    0.0500    0.0001
    0.0500    0.8999
```

```
CondHearGivenTiger=PCONJUNTA./marg_TLTR
```

```
CondHearGivenTiger = 2×2
    0.5000    0.0001
    0.5000    0.9999
```

```
sum(CondHearGivenTiger) %columnas suman 1
```

```
ans = 1×2
     1     1
```

No es una sorpresa el resultado, porque esos eran los datos iniciales.

La otra condicional es más curiosa:

```
CondTigerGivenHear=PCONJUNTA./marg_HLHR
```

```
CondTigerGivenHear = 2×2
    0.9982    0.0018
    0.0526    0.9474
```

```
sum(CondTigerGivenHear,2) %filas suman 1
```

```
ans = 2×1
    1.0000
    1.0000
```

La condicional "p(<lado tigre> | observación)" tiene sentido matemático (verifica definiciones), pero (a primera vista) no "conceptual" en mi problema. La observación es el "efecto" y el tigre la "causa" (condición)... el tigre se pone en un sitio y entonces se escucha el rugido dependiendo del sitio y no al revés: afirmar que "se oye el ruido y el tigre decide dónde se pone basado en el ruido" no tiene "sentido físico"...

Pero la fórmula ES CORRECTA, y es la base de la inferencia bayesiana... Aunque la "cadena causal" sea la otra (primero decide lado y luego ruge), la probabilidad conjunta es la que es, y yo puedo "calcular probabilidades de dónde está el tigre tras escuchar un rugido a la izquierda"... a esa "estimación" le denominamos "probabilidad a posteriori" o "creencia".

Esa es la idea CLAVE de la "fórmula de Bayes", en materiales posteriores a este.