Bayes formulae: "hidden tiger" case study

Presentations in video:

http://personales.upv.es/asala/YT/V/tiger6brEN.html

http://personales.upv.es/asala/YT/V/tiger7rbsEN.html

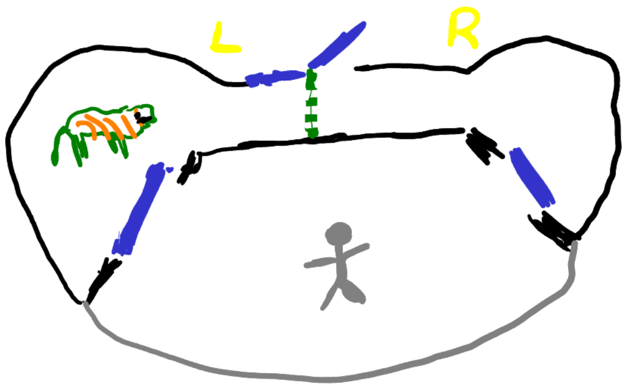
*This code executed in Matlab **`R2023b`**

Objectives: Illustrate the RECURSIVE Bayes formula in the case study "Tiger behind two doors" where several noisy measurements about tiger location need to be accumulated to improve chances of survival (correctly assessing where the tiger is before "acting" opening a door).

Table of Contents

Problem statement

A zoo keeper cleans the tiger attraction and, when finishing, he must decide which door to open to get out without being eaten by the tiger.



Tiger lies either left or right when central gate is closed, and roaring is heard either at the left or right doors... We'll assume "binary" setup:

Measurements (also denoted as "observations") are binary: `hear_left` (1 HL), `hear_right` (2 HR).

Tiger "states" are binary: `tiger_left` (1 TL), `tiger_right`(2 TR).

```
hear_L=1;  hear_R=2; %numbers indicate position on a list, variables are cathegorical and cannot be ordered/quantified.
tiger_L=1; tiger_R=2; %numbers indicate position on a list, variables are cathegorical and cannot be ordered/quantified.
```

Case 0: Central gate closes and I hear just ONE roar.

We'll define PM(m,s) is the probability of measuring "m" when state is "s". So, it is a CONDITIONAL probability table:

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

$$p(HR|TL) \quad p(HR|TR) \text{]}$$

Rows represent "observations", columns represent tiger location "states".

COLUMNS must add 1. PM does not need to be symmetric.

```
%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=[.99 .01;.01 .99];
%PM=[.55 .001;.45 .999];
PM
```

PM = 2×2

```
    0.9900    0.0100
    0.0100    0.9900
```

Note that this is a table of ***conditional*** probabilities, not a "joint" probability distribution.

Belief

Belief b is my "**prior**" probability of `tiger_right`. Of course, prior probability of `tiger_left` will be $1 - b$.

```
syms b real
```

Posterior belief update: Bayes rule

If I measure `hear_left`, in terms of conditional tables:

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

```
bnext_ifHL(b) = PM(hear_L,tiger_R)*b/
(PM(hear_L,tiger_R)*b+PM(hear_L,tiger_L)*(1-b));
vpa(bnext_ifHL)
```

ans(b) =

$$-\frac{0.01\,b}{0.98\,b - 0.99}$$

```
vpa(bnext_ifHL)
```

ans(b) =

$$-\frac{0.01\,b}{0.98\,b - 0.99}$$

If I measure `hear_right`,

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

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

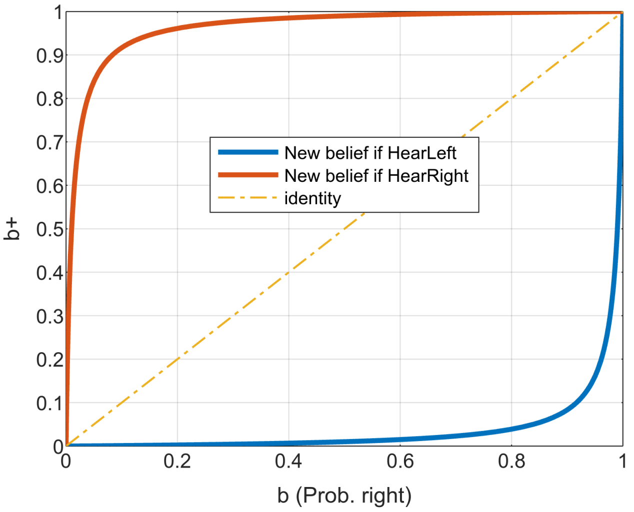
ans(b) =

$$\frac{0.99\,b}{0.98\,b + 0.01}$$

Graphical representation:

From a prior belief and an observation I get a posterior belief.

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



There is no "fixed" point where belief new will be equal to old except total belief on right (b=1) or left (b=0). If I am "100% sure" of something, experimenting will not provide any new information.

Review of Bayes formula

Let's use a compact notation for the conditional probability matrix:

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

HL HR (rows) given TL and TR (columns), respectively. If $b = p(TR)$, Bayes formula will say that, depending on whether I observe HL or HR, I'll apply one of these two belief updates:

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

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

Bayes formula for TWO roar observations (NON recursive)

If I wait until TWO roars are heard (in abstract notation, observing $o_1 \in \{HL, HR\}$ y $o_2 \in \{HL, HR\}$), assuming statistical independence (well, actually "conditional" independence because we are dealing with products of conditional probabilities), the new conditionals with respect to TL are:

$$p((HL, HL)|TL) = p(HL|TL) \cdot p(HL|TL) = q^2,$$

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

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

and the new conditionals conditioned to TR:

$$p((HL, HL)|TR) = r^2,$$

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

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

Thus, my new conditional probability table with two observations (well, abstracted to one observation with four possible outcomes) will be:

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

Note: meaning of $o_2 \wedge o_1$ refers to observations in a different instant of time: of course, we cannot observe $o_1 = HL,\ o_2 = HR$ if they referred to the same instant of time!. Problems with random variables in different time instants are said to study "dynamics" of a system.

Formally, we should denote $o_1 \in \{HL_{t=1}, HR_{t=1}\}$ and $o_2 \in \{HL_{t=2}, HR_{t=2}\}$, but temporal sub-indices will be removed for notational simplicity.

Bayes formula for the observation $(HL_{t=1}, HL_{t=2})$... shorthanded toHLL:

$$b_{++} = p(TR \mid HLL) = \frac{p(HLL \mid TR)p(TR)}{p(HLL \mid TL)p(TL) + p(HLL \mid TR)p(TR)} = \frac{r^2 \cdot b}{q^2 \cdot (1-b) + r^2 \cdot b}$$

We then should complete all options:

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

*Left as an exercise.

Probabilities $p(TL| \ldots)$ will be 1- $p(TR| \ldots)$... Thus, we don't need to compute/store them.

If we listen until hearing 3 roar events (assuming central gate closed, tiger cannot switch side), assuming "conditional independence", the table for later application of Bayes formula when actualk observations come would be:

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

Reduction to table where "only count matters"

If events are "counting 2L and 1R" (order does not matter), then the table can be summarised to:

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

The Bayes formula will end up giving the same result: dividing each element of the row associated with the actual observation by the sum of the row yields the same result whether or not it is multiplied by 3. So, observation ordering in this type of problems is usually disregarded, as the number of possible events is reduced and tables get smaller (imagine all the combinations with 20 roaring events!)

Bayes formula for an arbitrary number of roaring events (non-recursive)

With suitable combinatorics, we can summarise everything by saying that if, from a total of $m+n$ roaring events, we hear "m from the left side" and "n from the right side", we'll apply Bayes with:

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

Statistics textbooks (look for "***binomial***" distribution in your books) usually do it this way.

They will say that the conditional probabilities have a binomial distribution, specifically:

● the number of "left" roars out of a total of "Tot" roar observations (conditional on **TL**) is a random variable with binomial distribution of parameters (Tot, q).

● the number of "left" roars out of a total of "Tot" roar observations (conditional on **TR**) is a random variable with binomial distribution of parameters (Tot, r).

Numerical example:

Let us read the conditional probability table we started everything with:

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

```
q = 0.9900
```

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

```
r = 0.0100
```

If I hear "***m***" times roaring on the left and "***n***" on the right, then:

```
m=5; %number of HL events
n=3; %number of HR events
Tot=m+n; %total roaring events, número rugidos totales
p_LmRn_cond_TL = nchoosek(Tot,m) * q^m * (1-q)^n
```

```
p_LmRn_cond_TL = 5.3255e-05
```

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

```
p_LmRn_cond_TR = 5.4337e-09
```

So, applying Bayes, the "*a posteriori*" probability that the tiger is on the left or on the right will be:

```
b=0.5; %prior probability of 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 = 0.9999
```

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

```
p_TR_cond_LmRn = 1.0202e-04
```

Recursive Bayes formula

Can we recursively apply the formula with the original ONE roar table to obtain the same result for several roars? The answer is YES. Let's first look at the "two" roar case and then the general case.

This is the formula for a single roar event:

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

Conjecture: what will we get with this?

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

Let us substitute b_+, giving:

$$b_{++} = p(TL|LL) = \frac{r * \dfrac{r \cdot b}{q \cdot (1-b) + r \cdot b}}{q * \dfrac{q \cdot (1-b)}{q \cdot (1-b) + r \cdot b} + r * \dfrac{r \cdot b}{q \cdot (1-b) + r \cdot b}}$$

Then, multiplying numerator and denominator by $q \cdot (1-b) + r \cdot b$, that yields:

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

which coincides with the application of Bayes formula to the table of four possible observation outcomes PM2.

Let's discuss the general case.

Theorem: In general, if a set of past observations $O_N = \{o_1, o_2, \ldots, o_N\} \in \{HL, HR\}^N$ gave rise to a belief (posterior probability):

$$b_N = p(TR|O_N) = \frac{p(O_N|TR) \cdot b}{p(O_N|TL) \cdot (1-b) + p(O_N|TR) \cdot b}$$

then, given a new observation $o_{N+1} \in \{HL, HR\}$ the posterior probability is:

$$b_{N+1} = p(TR \mid O_{N+1}) = \frac{p(o_{N+1}|TR) \cdot b_N}{p(o_{N+1}|TL) \cdot (1 - b_N) + p(o_{N+1}|TR) \cdot b_N}$$

being $O_{N+1} = O_N \wedge o_{N+1} = \{o_1, o_2, \ldots, o_N, o_{N+1}\} \in \{HL, HR\}^{N+1}$, if the new observation is conditionally independent of the previous ones.

Proof: indeed, replacing the formula for b_N results in:

$$b_{N+1} = p(TR \mid O_N \wedge o_{N+1})$$

$$= \frac{p(o_{N+1}|TR) \cdot \left(\frac{p(O_N|TR) \cdot b}{p(O_N|TL) \cdot (1-b) + p(O_N|TR) \cdot b}\right)}{p(o_{N+1}|TL) \cdot \left(\frac{p(O_N|TL) \cdot (1-b)}{p(O_N|TL) \cdot (1-b) + p(O_N|TR) \cdot b}\right) + p(o_{N+1}|TR) \cdot \left(\frac{p(O_N|TR) \cdot b}{p(O_N|TL) \cdot (1-b) + p(\ldots}\right.}$$

so removing denominators:

$$b_{N+1} = \frac{p(o_{N+1}|TR)p(O_N|TR) \cdot b}{p(o_{N+1}|TL)p(O_N|TL) \cdot (1-b) + p(o_{N+1}|TR)p(O_N|TR) \cdot b}$$

Thus, if O_N y o_{N+1} are CONDITIONALLY INDEPENDENT, i.e.,

$p(o_{N+1} \mid TR)p(O_N \mid TR) = p(O_N \wedge o_{N+1} \mid TR) = p(O_{N+1} \mid TR)$, results in

$$b_{N+1} = \frac{p(O_{N+1} \mid TR) \cdot b}{p(O_{N+1} \mid TL) \cdot (1-b) + p(O_{N+1} \mid TR) \cdot b}$$

i.e., Bayes formula with the probability of the joint event (old observations plus new one).

In summary, for "conditionally independent" observations, we can apply Bayes formula with an "expanded table" or just apply it recursively as new observations are gathered.

Practical usefulness of recursive Bayes formulae:

In some problems we do not know a priori how many data points (roars, observations, experiments) we need... We will accumulate data "until we have enough evidence"... If the tiger can change places then the complete table would have 2^10 observations (rows) and 2^10 possible sequences of tiger positions (columns)... Everything becomes very complicated, while in the recursive formula we only deal with "a single number that is updated with each observation."

This is the main idea being "recursive Bayes filters", topic of other materials; "Bayesian networks" also root on being able to explain a complex statistical model as the interaction between several conditionally independent variables.

Simulation of Recursive Bayes filter (tiger does NOT switch side)

```
Posterior_ifL=matlabFunction(bnext_ifHL); %symbolic to numeric
Posterior_ifR=matlabFunction(bnext_ifHR); %symbolic to numeric

real_tiger=tiger_R; %actual location of tiger

belief_ini=0.5; %initial belief, non-informative prior
N=15; %number of samples per simulation, first sample is "prior",
so N=observations+1
Nsimulations=40;
belief_plot=zeros(N,Nsimulations);
observ_plot=zeros(N,Nsimulations);
t_fin=zeros(Nsimulations,1);

for Nsim=1:Nsimulations
endflag=false;
belief=belief_ini;
belief_plot(1,Nsim)=belief_ini;
observ_plot(1,Nsim)=NaN; %prior does not have observation

%simulate, if so wished, a given sequence of observations:
%obs_test=[hear_L hear_L hear_R hear_L  hear_R hear_L hear_R
hear_L];

    for i=2:N
        % Measurement simulation (uncertain)
        % Simulate Bernouilli (unfair coin) distribution from
uniform
        if rand()<PM(hear_L,real_tiger) %rand returns uniform in
[0,1]
            obs=hear_L;
        else
            obs=hear_R;
        end

        %obs=obs_test(i-1); %If I wish to simulate a given sequence
of
        % observations instead of generating them randomly

        % Recursive Bayes Filter start
        if obs==hear_L %hear left
```

```
            belief=Posterior_ifL(belief); %belief update based on
measuring LEFT
        else %hear right
            belief=Posterior_ifR(belief); %belief update based on
measuring RIGHT
        end
        % Recursive Bayes Filter END

        %check termination time if N is very large
        if abs(belief-(real_tiger-1))<1/5000 && endflag==false
            t_fin(Nsim)=i;
            endflag=true;
        end
        observ_plot(i,Nsim)=obs; %store observation data
        belief_plot(i,Nsim)=belief; %store posterior probability
(belief) data
    end
end
```

The belief on tiger being at the right-hand side posterior to all the accumulated evidence is:

```
belief
```

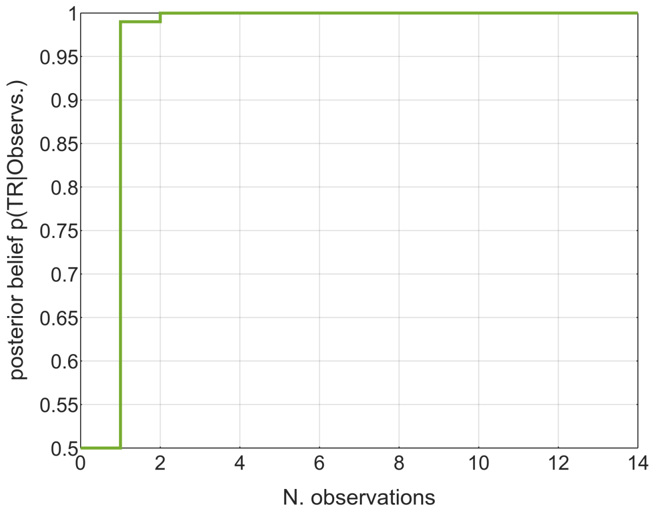
```
belief = 1
```

Time to be 4999 out of 5000 sure tiger is at a given side:

```
[min(t_fin) mean(t_fin) max(t_fin)]
```

```
ans = 1×3
     3     3     3
```

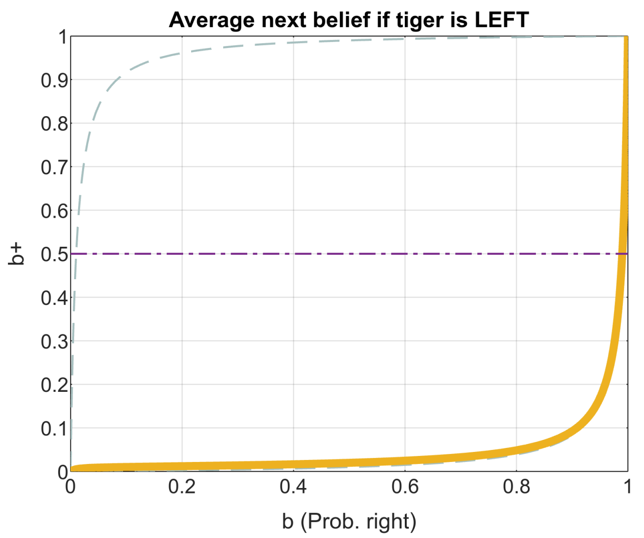
```
stairs(0:(N-1),belief_plot,LineWidth=1.5), grid on
if Nsimulations==1
    hold on
    plot(0:(N-1),observ_plot-1,'.',MarkerSize=14)
    xline(t_fin,'-.')
    hold off
end
hold off, ylabel("posterior belief p(TR|Observs.)"),xlabel("N.
observations")
```



If tiger truly were TL, then on average, next belief would be:

Si el tigre "de verdad" está a la izquierda, entonces en media, la siguiente creencia será:

```
fplot([bnext_ifHL;bnext_ifHR],[0 1],'--',LineWidth=1,Color=[.65 .75
.75]),hold on
fplot([bnext_ifHL*PM(hear_L,tiger_L)+bnext_ifHR*PM(hear_R,tiger_L)],
[0 1],LineWidth=4), grid on
hold on, fplot(b,[0 1],'-.',LineWidth=1), hold off
xlabel('b (Prob. right)'), ylabel('b+'), title("Average next belief
if tiger is LEFT")
```



If tiger truly were TR, then on average, next belief would be:

```
fplot([bnext_ifHL;bnext_ifHR],[0 1],'--',LineWidth=1,Color=[.65 .75
.75]),hold on
fplot([bnext_ifHL*PM(hear_L,tiger_R)+bnext_ifHR*PM(hear_R,tiger_R)],
[0 1],LineWidth=3), grid on
 fplot(b,[0 1],'-.',LineWidth=1), hold off
```

```
xlabel('b (Prob. right)'), ylabel('b+'), title("Average next belief
if tiger is RIGHT")
```