Primer: conditional probabilities

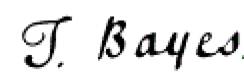
Lecture 11.1

Mathematical predictions

- We can ‘predict’ where the spacecraft will be at noon in 2 months from now
- We cannot predict where you will be tomorrow at noon
- But, based on numerous observations, we can estimate the probability

Bayesian beliefs

- How do we judge that something is true?
- Can mathematics help make judgments more accurate?
- Bayes: our believes should be updated as new evidences become available



Bayes' method

- There are 2 events: **A** and not A (**B**) which you believe occur with probabilities P(**A**) and P(**B**). Estimation P(**A**):P(**B**) represents odds of A vs. B.
- Collect evidence data **E.**
- Re-estimate P(**A**|**E**):P(**B**|**E**) and update your beliefs.

Example (fictitious): hit-and-run

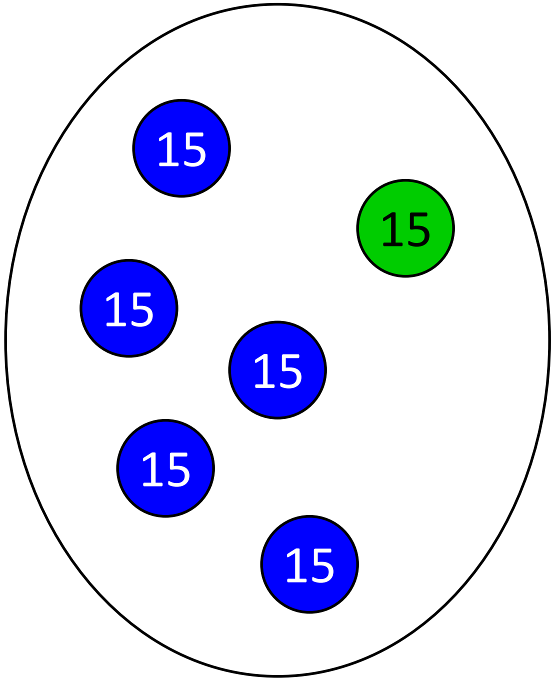
- 75 blue cabs (**B**) and 15 green cabs (**G**)
- P(**B**):P(**G**)=5:1

- At night: hit-and-run episode
- Witness: “I saw a green cab”: X_G
- Witness is tested at night conditions: identifies correct color 4 times out of 5

- Question: what is more probable:

B or **G**

?



Adopted from: **The numbers behind NUMB3RS:** solving crime with mathematics by Devlin and Lorden.

Probability

- Basic element: random variable

  e.g., *Car* is one of *<blue, ¬blue(green)>*

  Weather is one of *<sunny,rainy,cloudy,snow>*

- Both *Car* and *Weather* are discrete random variables
  - Domain values must be
    - exhaustive (blue and green – are all the cabs)
    - mutually exclusive (green is always not blue, there are no cars which are half green, half blue)

- Elementary propositions are constructed by the assignment of a value to a random variable:

  e.g., *Car = ¬blue*,

  Weather = sunny

Conditional probability

- P(A|B) – probability of event A given that event B has happened
- In our case we want to compare:
- the car was **G** given a witness testimony X_G: P(**G**|$\mathbf{X_G}$)
- vs.
- the car was **B** given a witness testimony X_G: P(**B**|$\mathbf{X_G}$)

Prior probability and distribution

- Prior or unconditional probability associated with a proposition is the degree of belief accorded to it in the absence of any other information.

  e.g.,

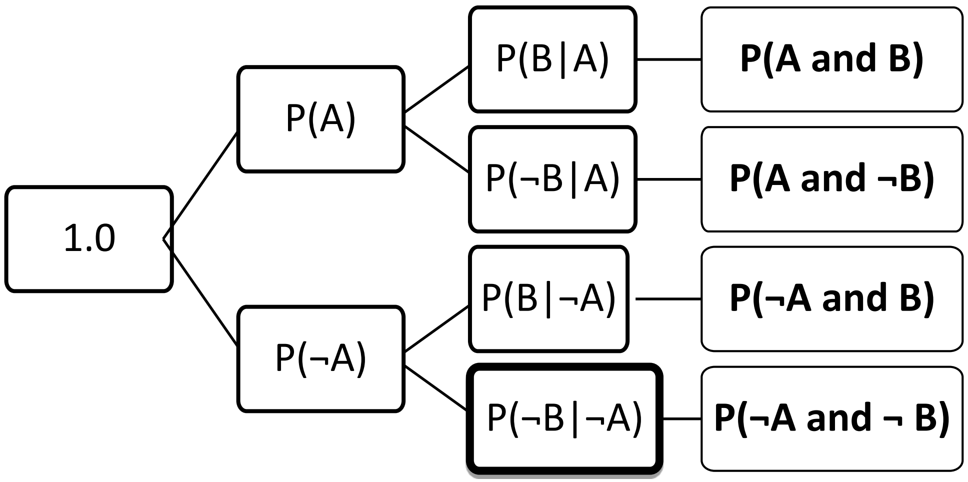
  P(Car = blue) = 0.83 (or abbrev. P(blue) = 0.83)
  P(Weather = sunny) = 0.7 (or abbrev. P(sunny) = 0.7)

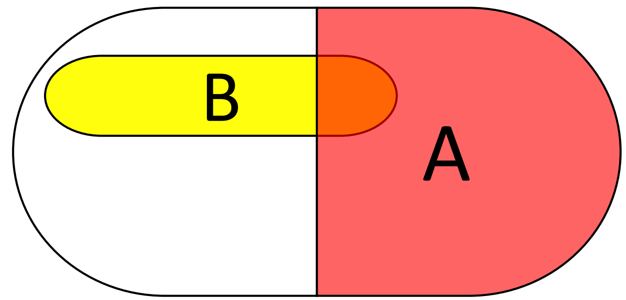
- Probability distribution gives probabilities of all possible value assignments:

  P(Weather = sunny) = 0.7
  P(Weather = rain) = 0.2
  P(Weather = cloudy) = 0.08
  P(Weather = snow) = 0.02

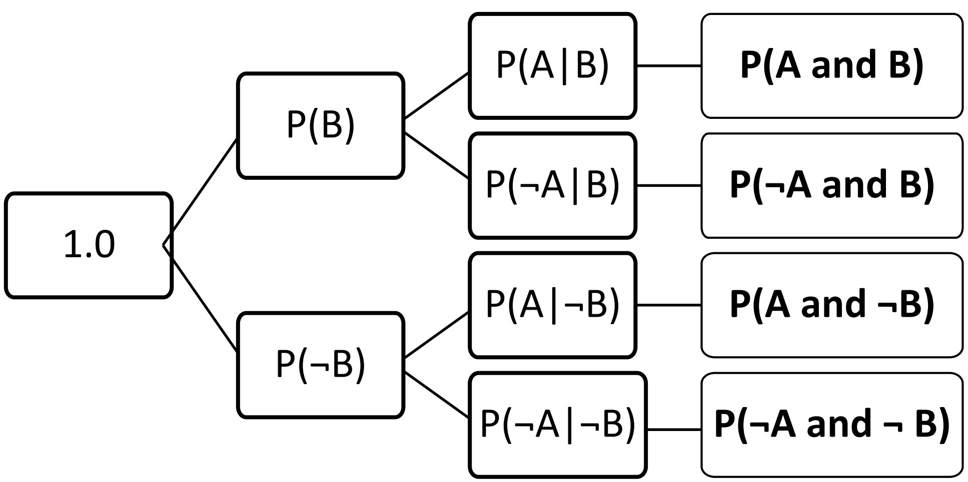
  - Sums up to 1.0

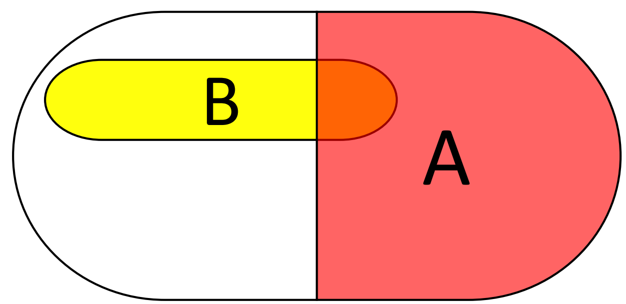
Two random events (not independent) happen at the same time – P(A and B)





Possible event combinations when we know the outcome of event A:
P(B|A)=1/12 and P(A)=1/2

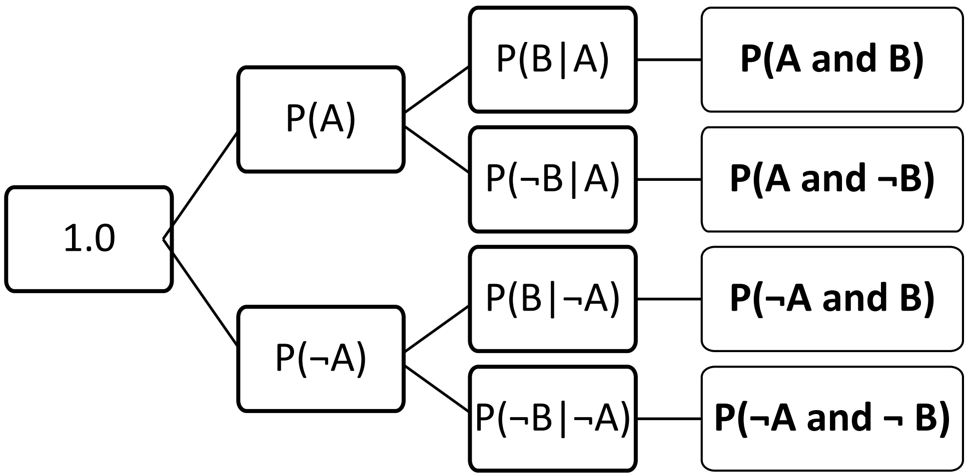




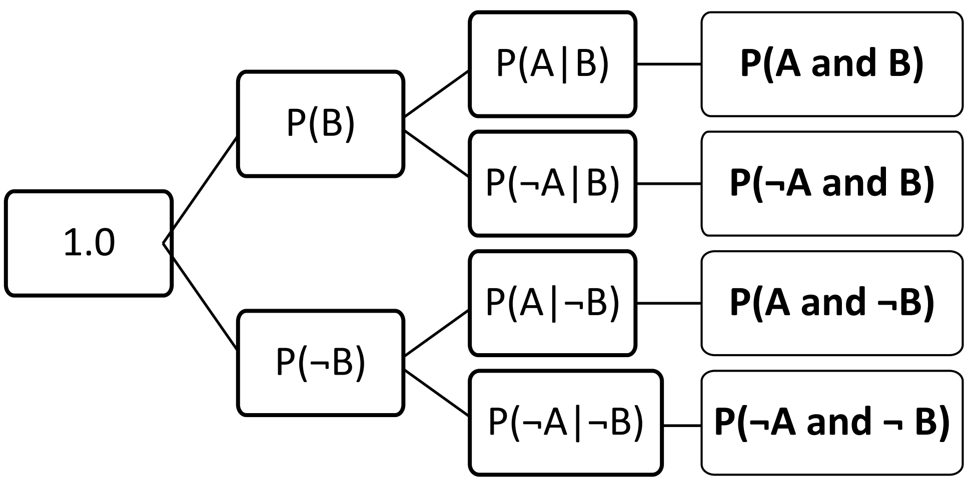
Possible event combinations when we know the outcome of event B:
P(A|B)=1/4 and P(B)=1/6

But in both cases P(A and B) is the same: orange area in the diagram

Intuition for Bayes's theorem



→ P(A and B)=P(A)*P(B|A)

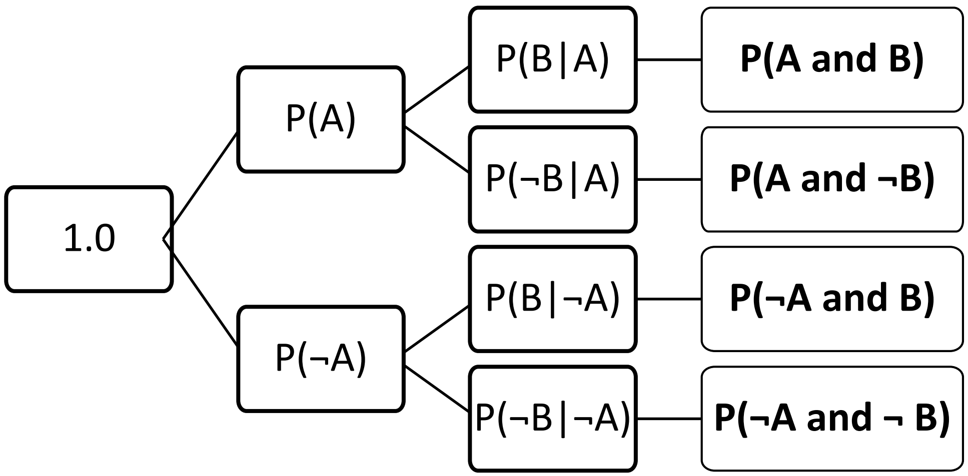


→ P(A and B)=P(B)*P(A|B)

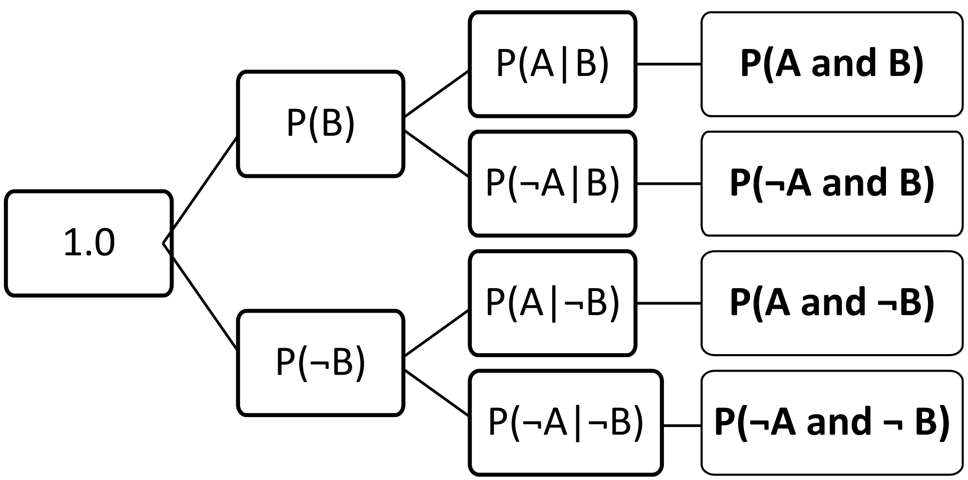
P(A and B)=P(A)*P(B|A)=P(B)*P(A|B)

P(¬A and B)=P(¬A)*P(B|¬A)=P(B)*P(¬A|B)

Bayes' theorem



→ P(A and B)=P(A)*P(B|A)

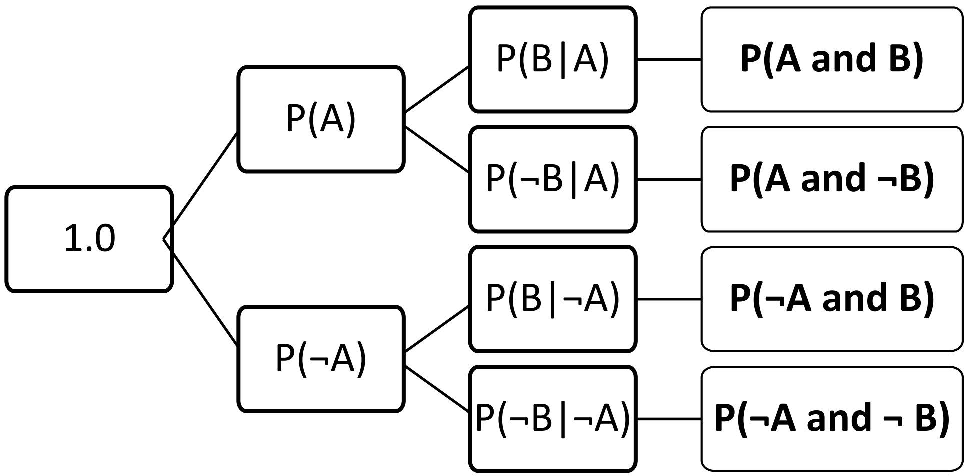


→ P(A and B)=P(B)*P(A|B)

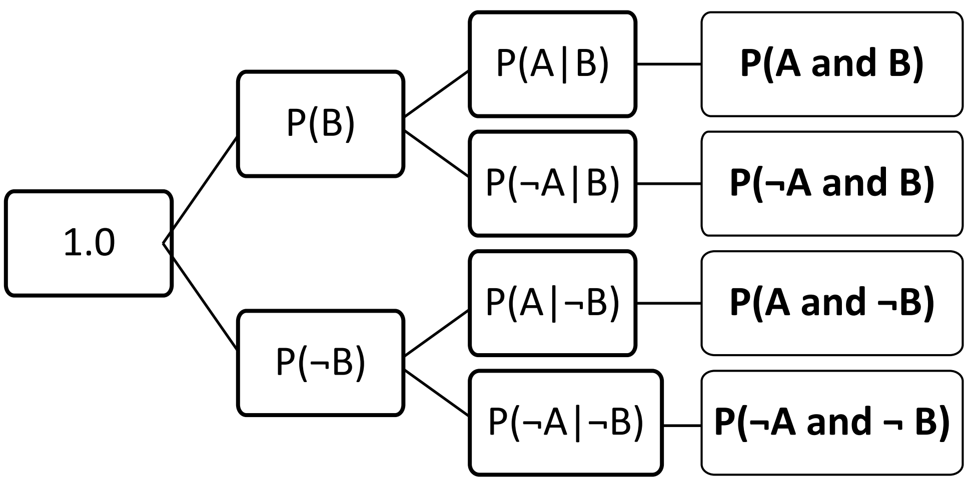
P(A)*P(B|A)=P(B)*P(A|B)

P(¬A)*P(B|¬A)=P(B)*P(¬A|B)

In other words:



→ P(A and B)=P(A)*P(B|A)



→ P(A and B)=P(B)*P(A|B)

P(A|B)=P(A)*P(B|A)/P(B)

P(¬A|B)=P(¬A)*P(B|¬A)/P(B)

Bayes' Rule for updating beliefs

$P(A|B)=P(A)*P(B|A)/P(B)$

$P(\neg A|B)=P(\neg A)*P(B|\neg A)/P(B)$

- We want to compare $P(A|B)$ and $P(\neg A|B)$, i.e. given evidence B what probability is higher: that A occurred or that ¬A occurred?

- We know $P(A)$ and $P(\neg A)$ – prior probabilities
- We know $P(B|A)$ and $P(B|\neg A)$

- From Bayes' theorem:

$$P(A|B) = P(A)*P(B|A) / P(B)$$

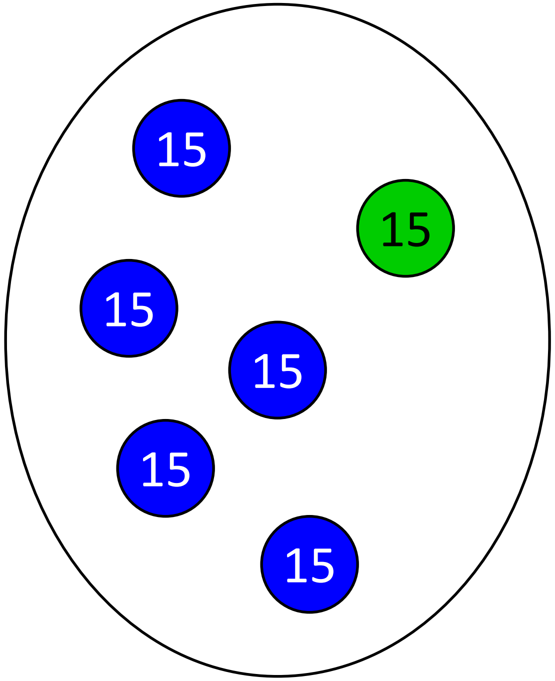
$$P(\neg A|B) = P(\neg A)*P(B|\neg A) / P(B)$$

Back to hit-and-run

- What is more probable: **B** or **G** ?

- All cabs were on the streets:
- Prior probabilities: P(**B**) =5/6, P(**G**) = 1/6

- The eyewitness test has shown:
- $P(X_G \mid G) = 4/5$ (correctly identified)
- $P(X_G \mid B) = 1/5$ (incorrectly identified)



$$\mathbf{P(G|X_G)} = P(G)*P(X_G|G) / P(X_G)$$

$$\mathbf{P(\neg G|X_G)} = P(\neg G)*P(X_G|\neg G)/P(X_G)$$

Bayes rule

Hit-and-run: solution

- $P(B) = 5/6$, $P(G) = 1/6$
- $P(X_G \mid G) = 4/5$, $P(X_G \mid B) = 1/5$

- Probability that car was **green** given the evidence X_G:
- $P(G|X_G) = P(G) * P(X_G|G) / P(X_G) = [1/6 * 4/5] / P(X_G) \quad = 4/30P(X_G)$
- //- 4 parts of $30P(X_G)$

- Probability that car was **blue** given the evidence X_G:
- $P(B|X_G) = P(B) * P(X_G|B) / P(X_G) = [5/6 * 1/5] / P(X_G) \quad = 6/30P(X_G)$
- //- 6 parts of $30P(X_G)$

- 6:4 odds that the car was **B**!

The probabilistic conclusion reached without knowing $P(X_G)$
– the probability of an actual event

Probabilistic predictions

Name	Toothache	…	Cavity
Smith	true	…	true
Mike	true	…	true
Mary	false	…	true
Quincy	true	…	false
…	…	…	…

- Given the evidence (data),

can we certainly derive

the diagnostic rule:

if Toothache=true then Cavity=true ?

- This rule isn't right always.
  - Not all patients with toothache have cavities - some of them have gum disease, an abscess, etc.

- We could try an inverted rule:

if Cavity=true then Toothache=true

- But this rule isn't necessarily right either - not all cavities cause pain.

Certainty and Probability

- The connection between toothaches and cavities is not a certain logical predicate in either direction.
- However, we can provide a **probability** that given an evidence (toothache) the patient has cavity.
- For this we need to know:
  - Prior probability of having cavity: how many times dentist patients had cavities: P(cavity)
  - The number of times that the evidence (toothache) was observed among all cavity patients: P(toothache|cavity)

Bayes' Rule for diagnostic probability

Bayes' rule:

$$P(A|B)=P(A)*P(B|A)/P(B)$$

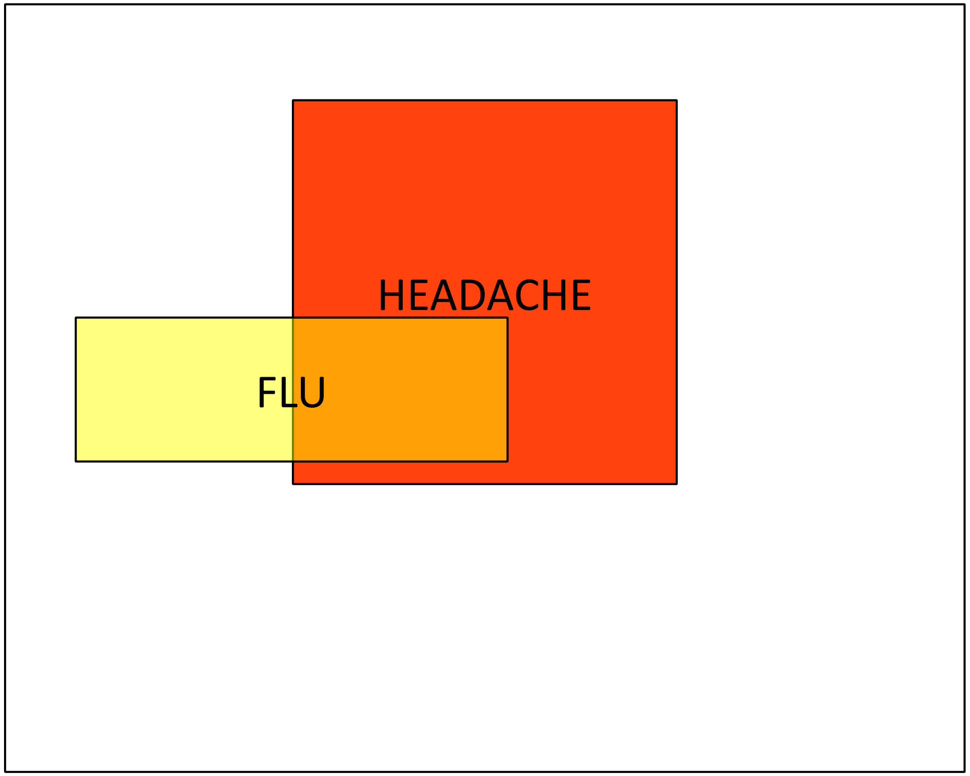
- Useful for assessing **diagnostic** probability from **symptomatic** probability as:
- P(Cause|Symptom) = P(Symptom|Cause) P(Cause) / P(Symptom)

- Bayes's rule is useful in practice because there are many cases where we do have good probability estimates for these three numbers and need to compute the fourth.

Bayes rule application. Example 1

$P(H)=1/10$
$P(F)=1/40$
$P(H|F)=1/2$

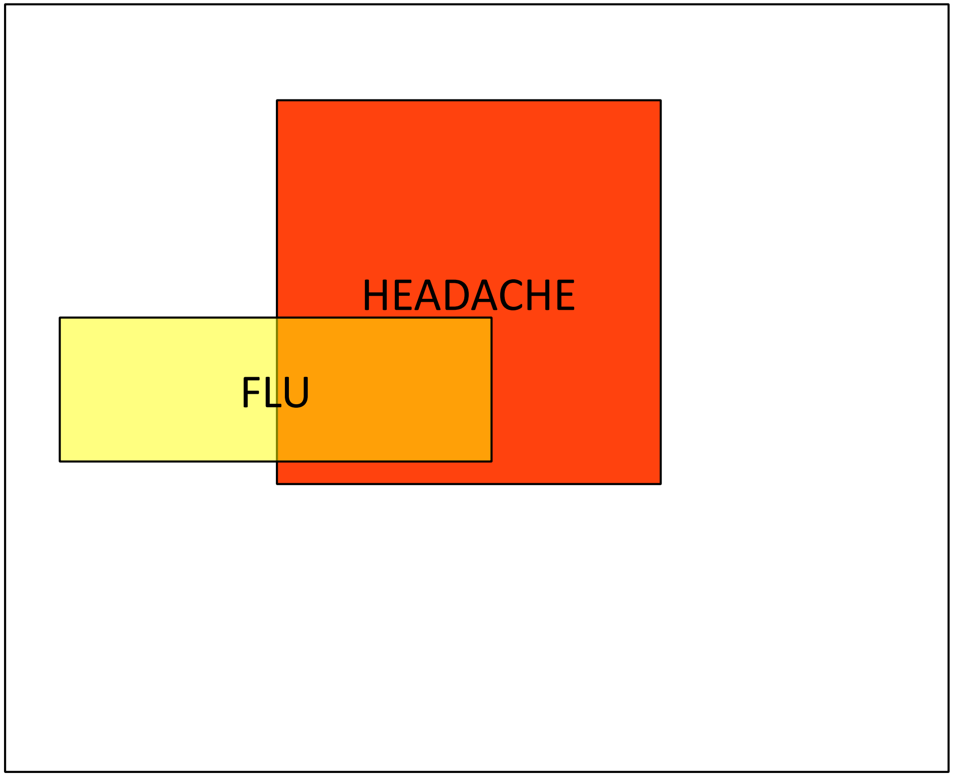
$P(F|H) = ?$



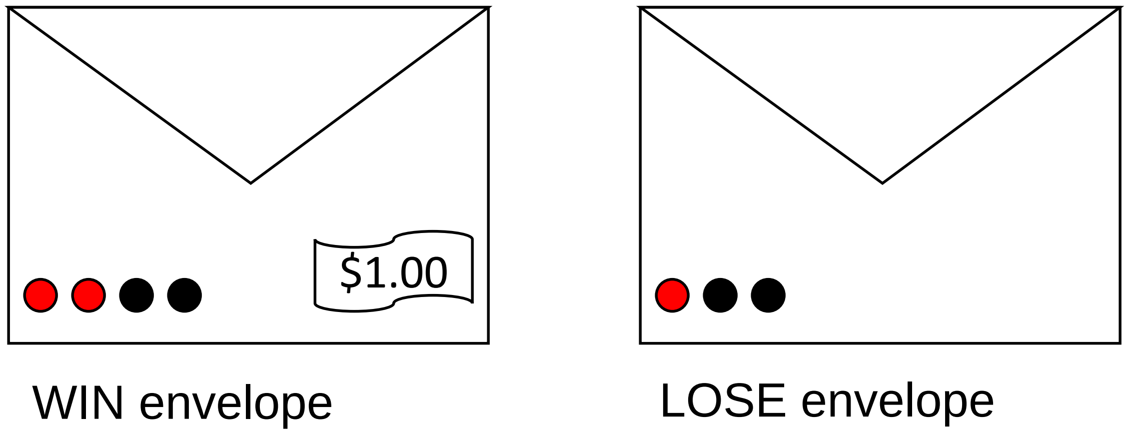
Bayes rule application. Example 1

P(H)=1/10
P(F)=1/40
P(H|F)=1/2

P(F|H) =P(H|F)P(F)/P(H)
=1/2*1/40 *10=1/8

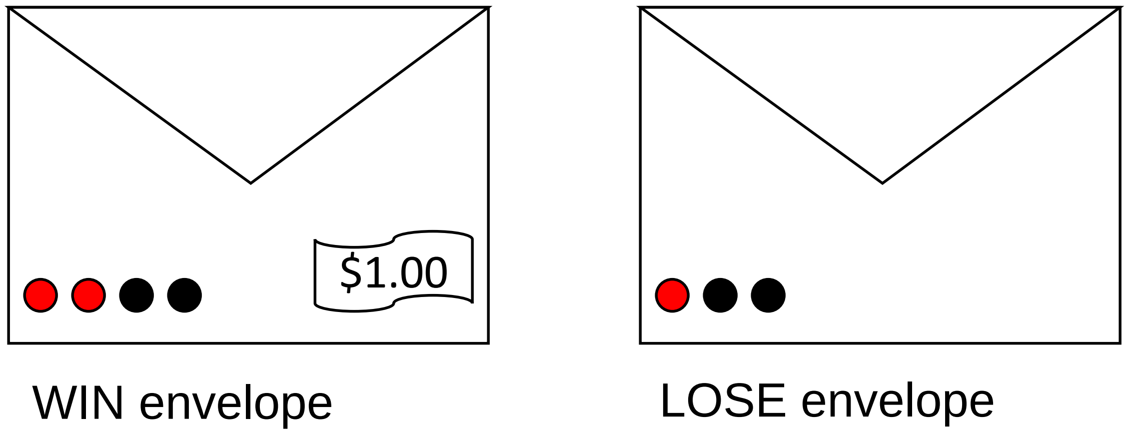


Bayes rule application. Example 2



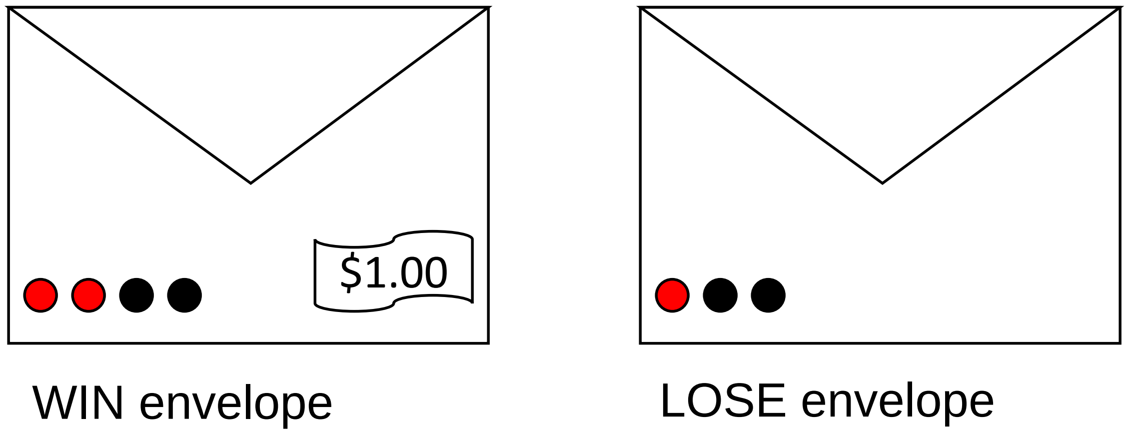
Someone draws an envelope at random and offers to sell it to you.
How much should you pay?
The probability to win is 1:1. Pay no more than 50c.

Bayes rule application. Example 2



Variant: before deciding, you are allowed to see one bead drawn from the envelope.
Suppose it’s black: How much should you pay?
Suppose it’s red: How much should you pay?

Bayes rule application. Example 2



Variant: before deciding, you are allowed to see one bead drawn from the envelope.

Suppose it's black: How much should you pay?

P(W|b)=P(b|W)P(W)/P(b) =(1/2*1/2)/P(b)=1/4 *1/P(b)

P(L|b)=P(b|L)P(L)/P(b)=(2/3*1/2)/P(b) = 1/3 * 1/P(b)

Probability to win is now 3:4 – pay not more than $(3/7)

Suppose it's red: How much should you pay? – the same logic

Log-odds ratio

- Note, that we do not have to know P(b) in order to make predictions: we just find the ratio of 2 mutually exclusive probabilities

$$\frac{P(W|b)=P(b|W)P(W)/P(b)}{P(L|b)=P(b|L)P(L)/P(b)}$$

- Instead of finding ratio, find its log:

$$\log \frac{P(W|b)=P(b|W)P(W)/P(b)}{P(L|b)=P(b|L)P(L)/P(b)}$$

If positive, then winning is more probable, if negative – loosing is more probable