Primer: conditional probabilities

Lecture 7.1

by Marina Barsky

BOOLEAN-VALUED RANDOM VARIABLES

Discrete Boolean random variables

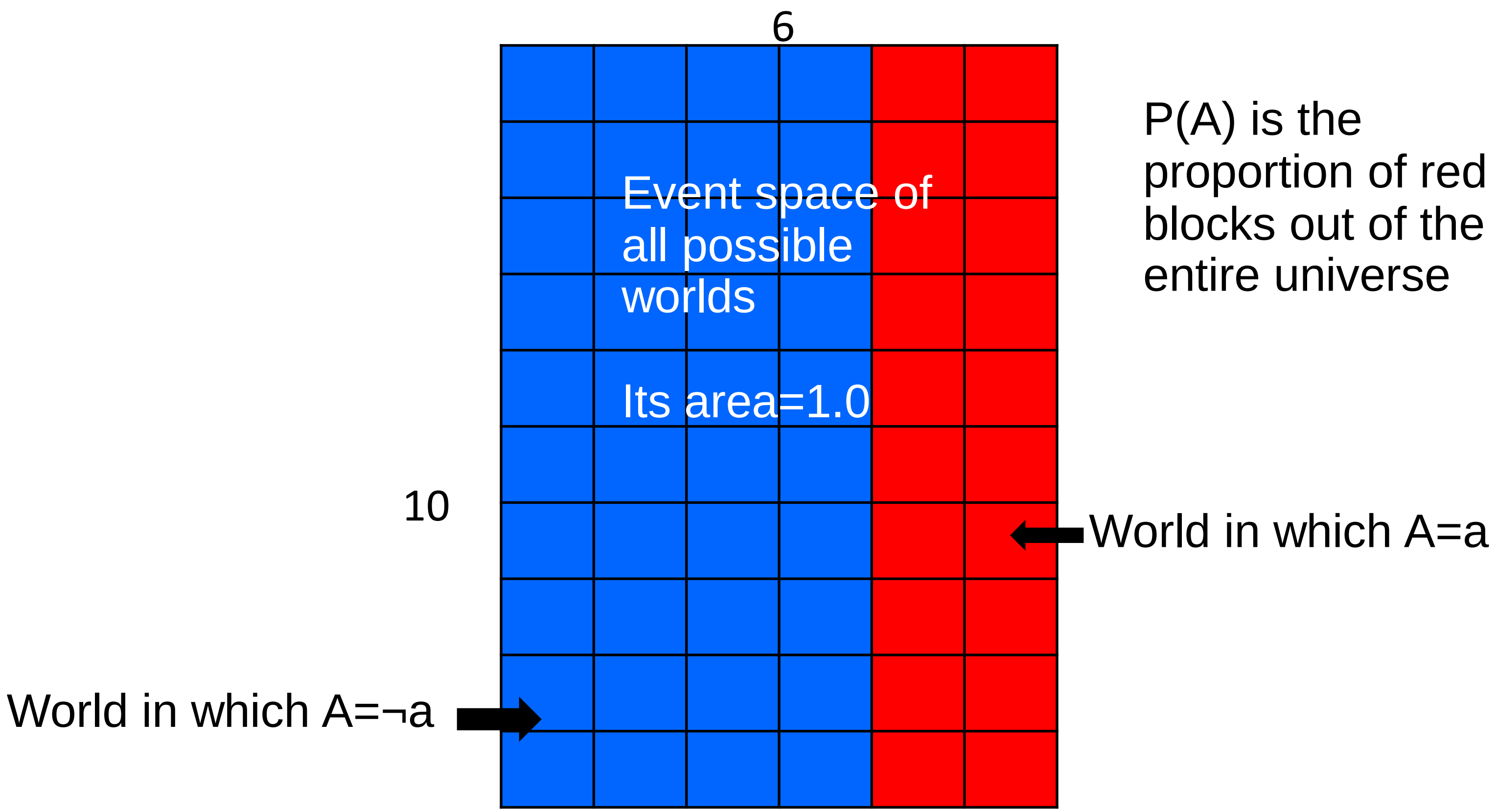
A is a Boolean-valued random variable if *A* denotes an event, and there is some degree of uncertainty as to whether *A* occurs or not.

Examples:

- P = True: The US president in 2023 will be male
- P=¬True: The US president will not be a male
- H = True: You wake up tomorrow with a headache
- H=¬True: No headache

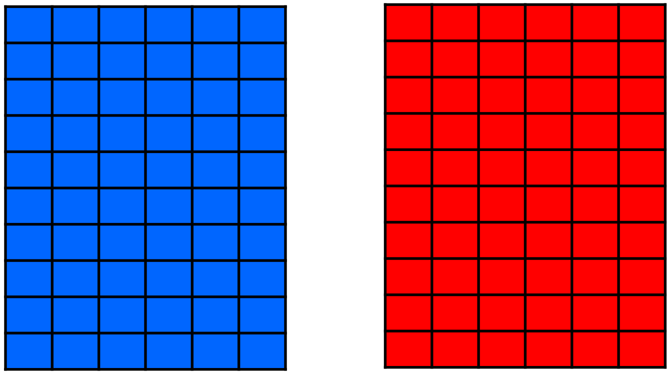
Discrete probabilities

We write P(A=a), or P(A=true) or simple P(A) as “the fraction of possible worlds where A is true”



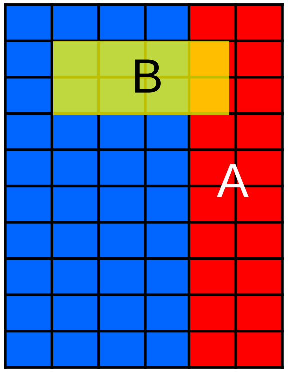
P(A) is the proportion of red blocks out of the entire universe

Axioms



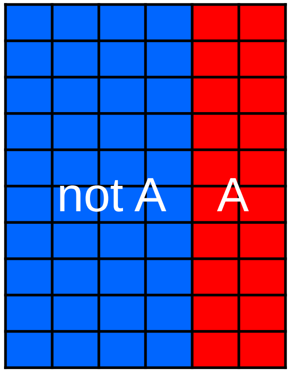
I. 0<= P(A=a)<=1.0

II. P(A or B)=P(A)+P(B)-P(A and B)



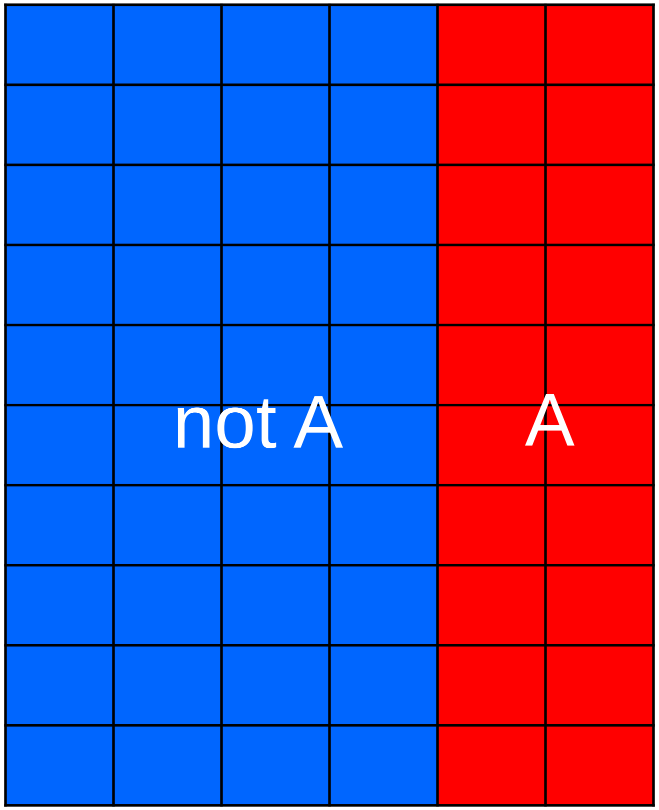
A B

III. P(A)+P(¬A)=1.0



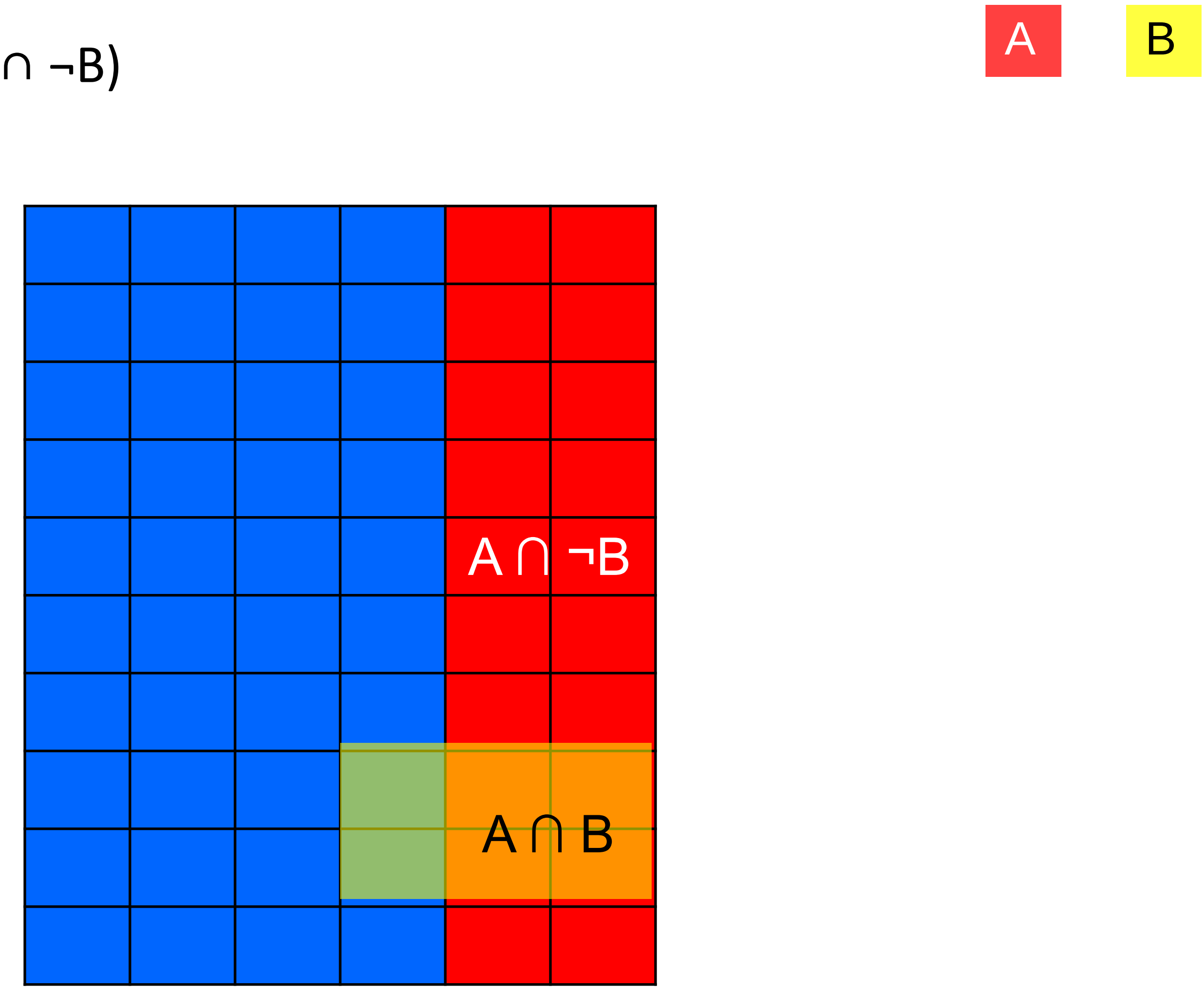
Theorem 1

P(¬A)=1-P(A)



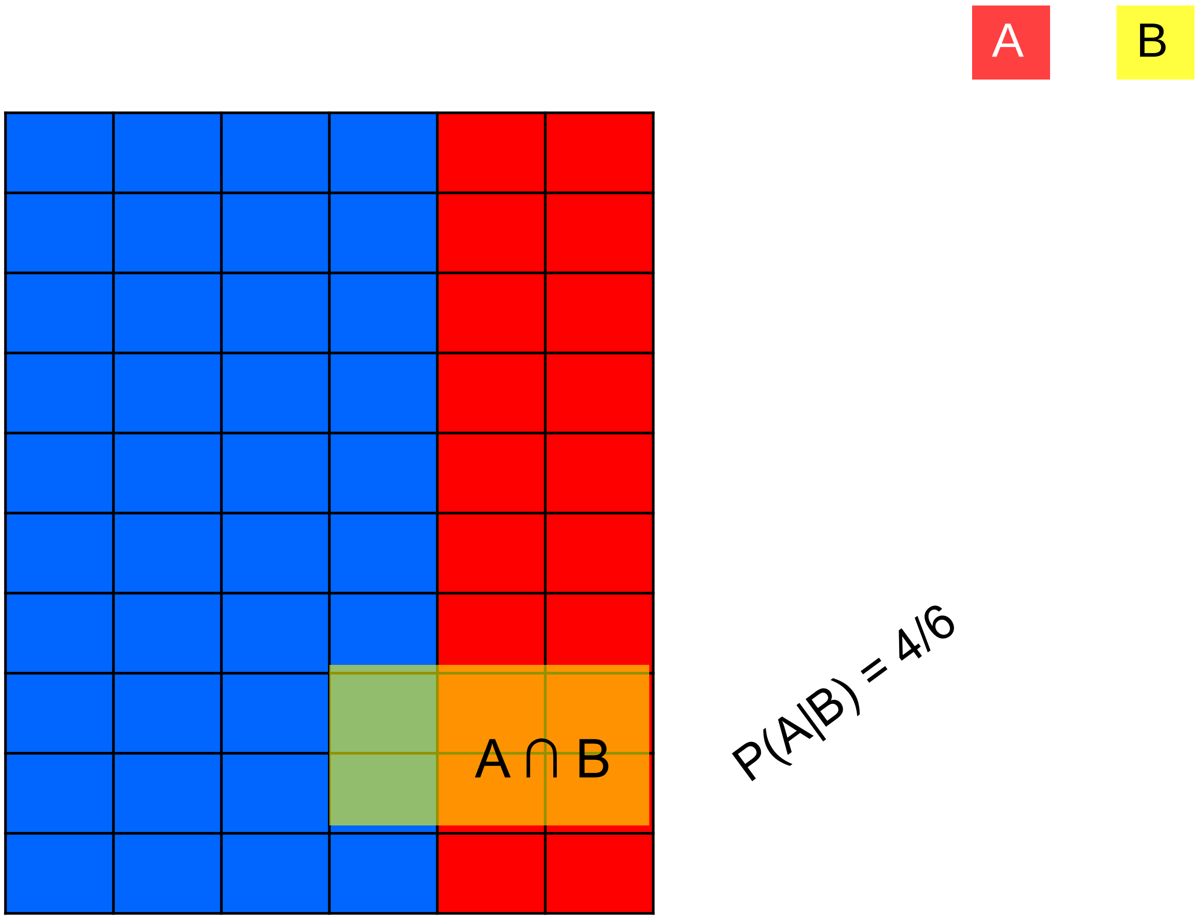
Theorem 2

$$P(A)=P(A \cap B) + P(A \cap \neg B)$$



Conditional probability: definition

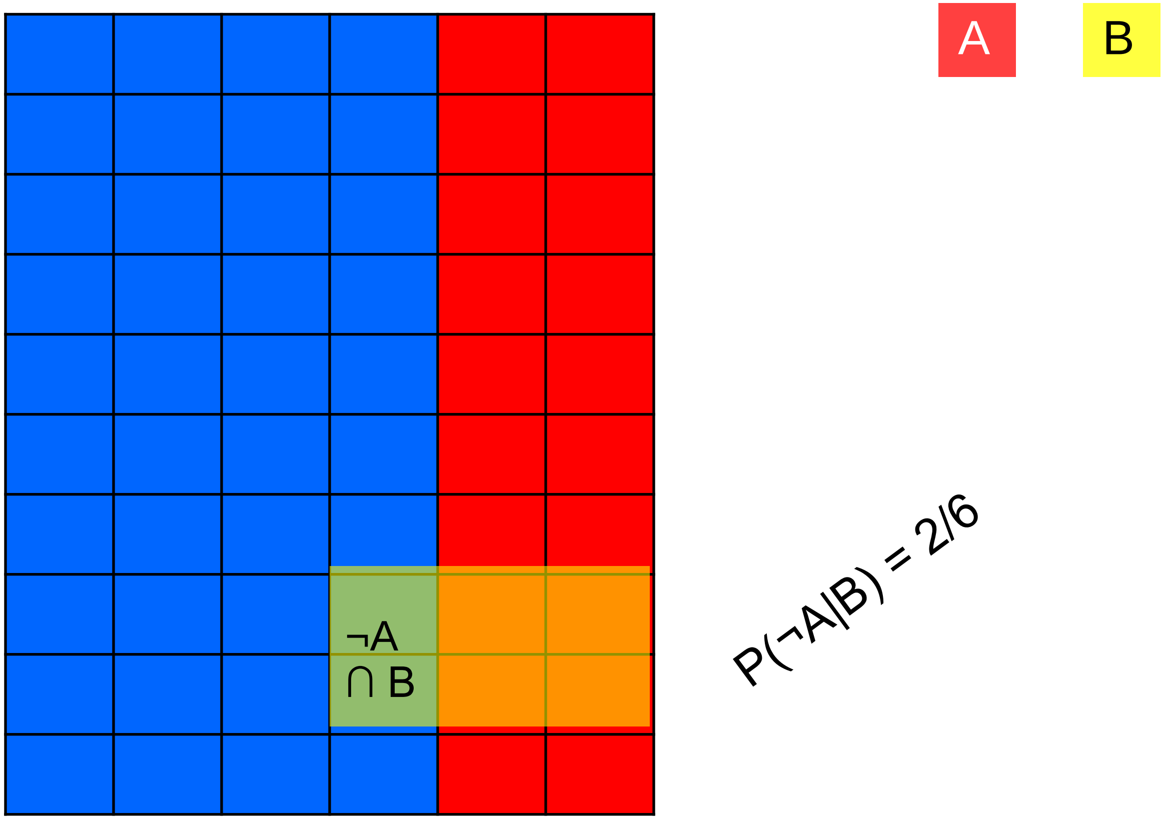
- P(A|B) = fraction of worlds in which A is true out of all the worlds where B is true



CP definition: $P(A|B) = P(A \cap B) / P(B)$

Conditional probability: definition

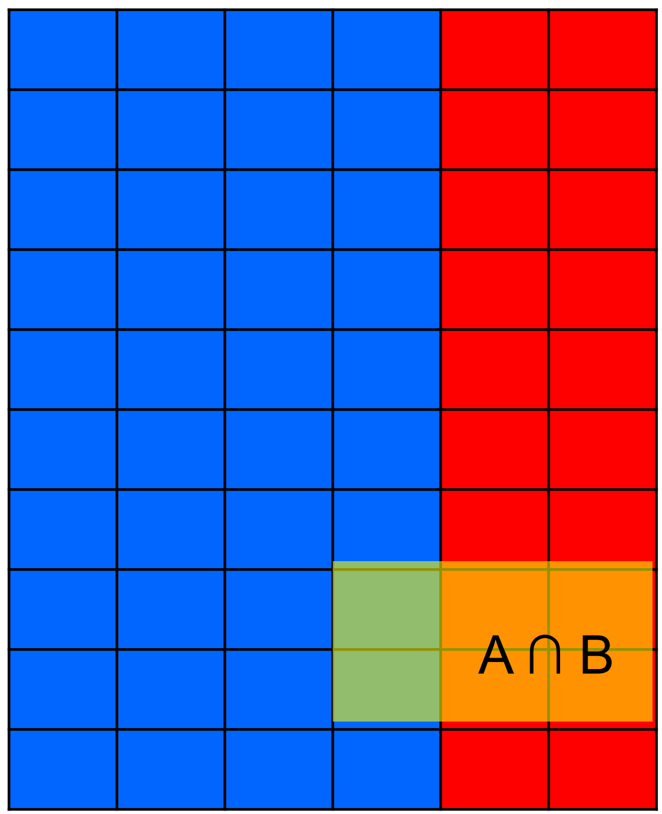
- P(A|B) = fraction of worlds in which A is true out of all the worlds where B is true



$$P\,(\neg A|B) = P(\neg A \cap B)\; /\; P(B)$$

Conditional probability: definition

- P(B|A) = fraction of worlds in which B is true out of all the worlds where A is true



A B

P(B) = 6/60 = 0.1 (unconditional)
P(B|A) = 4/20 =0.2 (conditional)

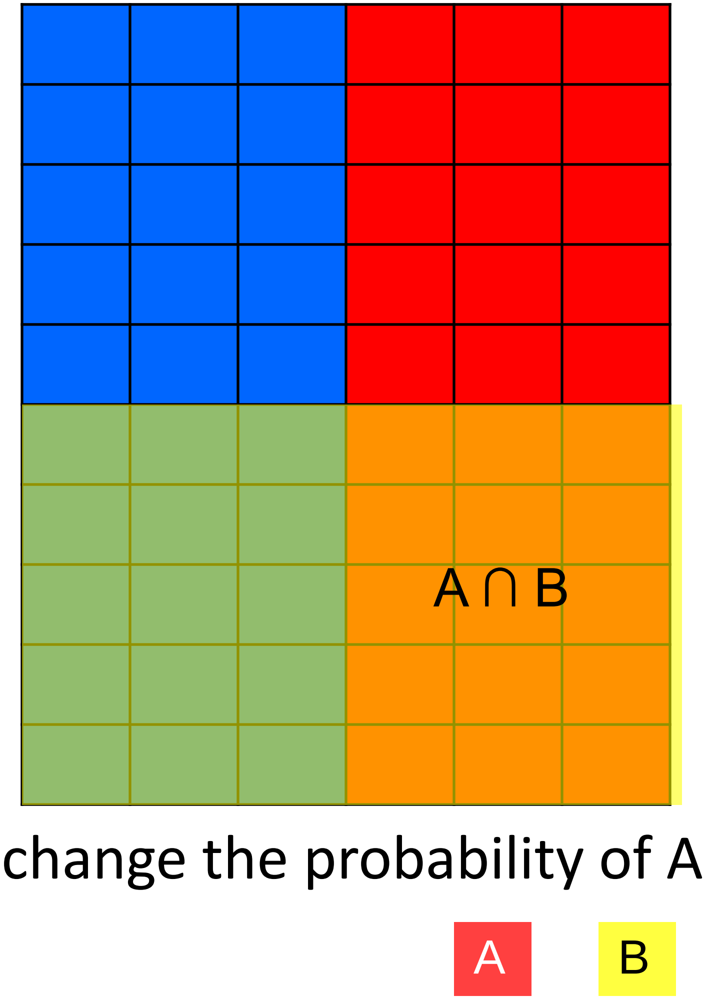
$P(A \cap B) = 4/60$
$P(A) = 20/60$
$P(B|A) = 4/60 : 20/60 = 0.2$

$$P(B|A)= P(A \cap B) / P(A)$$

Probabilistic independence

Two random variables A and B are called *mutually independent* if
P(A|B) = P(A):

P(A|B) = P(A)	15/30 = 30/60
P(¬ A|B) = P(¬ A)	15/30 = 30/60
P(A| ¬ B) = P(A)	15/30 = 30/60
P(¬A| ¬ B) = P(A)	15/30 = 30/60

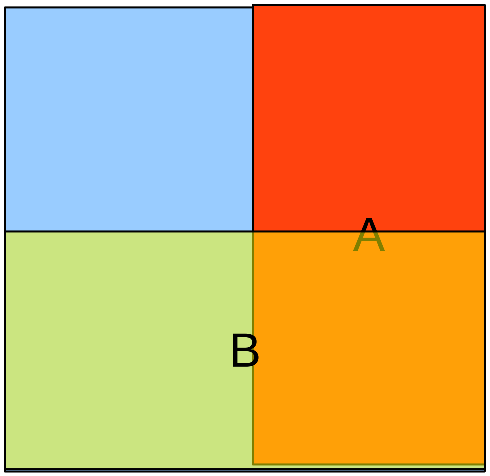


Knowing that B is true (or false) does not change the probability of A

Independent and mutually exclusive events

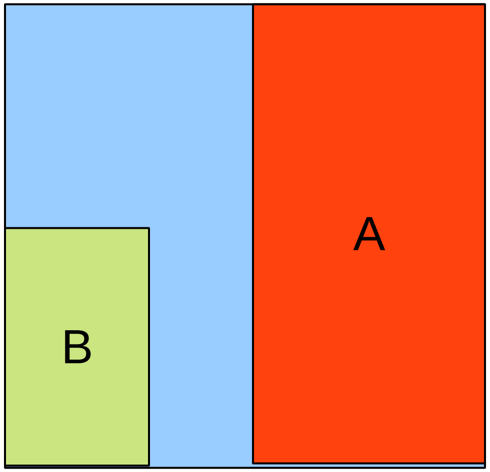
A is *independent* of B: knowing that B is true (or false) does not change the probability of A:

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



A and B are *mutually exclusive* – not independent variables: if A is true then B is false, if A is false then B is true with probability $P(B|\neg A)$

$$P(A \cap B)=0$$



Probability of two independent events

From the definition of conditional probabilities:

$$P(A|B)= P(A \cap B) / P(B)$$

we can compute $P(A \cap B)$ – that both events happened together:

$$P(A \cap B) =P(A|B)P(B)$$

If A and B are *independent* that becomes:

$$P(A \cap B) =P(A)P(B)$$

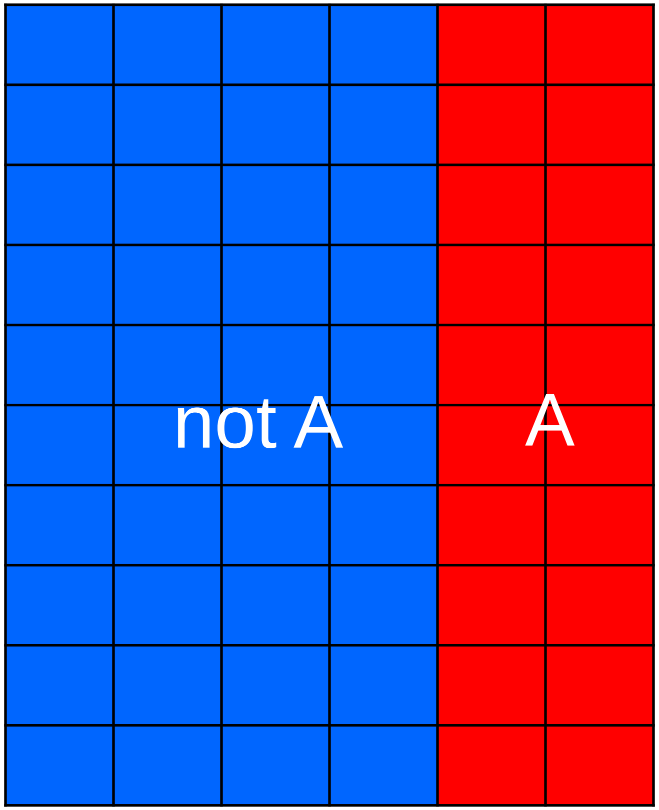
Probability of mutually exclusive events

A and ¬A are mutually exclusive, so by Axiom II:

P(A or B)=P(A)+P(B)-~~P(A and B)~~

And for A or ¬A :

P(A or ¬A)=P(A) + P(¬A) = 1



Probabilistic inductive reasoning

- Critical thinking: always have good reasons for your beliefs
- Some reasons may be 100% true - some only probable

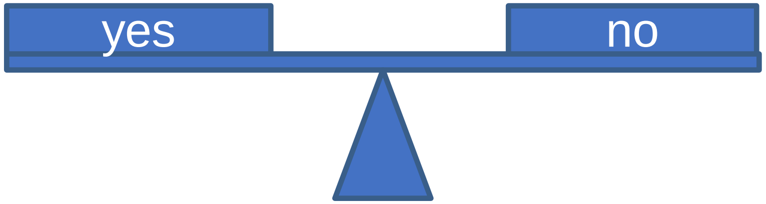
- Inductive reasoning with probabilities: you always have a chance of being wrong

http://www.starwars.com/video/never-tell-me-the-odds

I believe that John will not be at the party

In the absence of facts

John will **not** be at the party

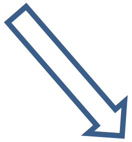


What are the odds?

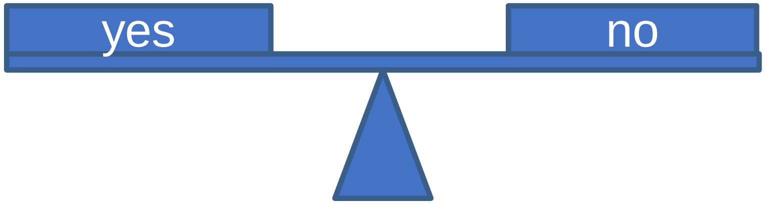
I believe that John will not be at the party

Invalid (illogical) reasoning

I do not like John



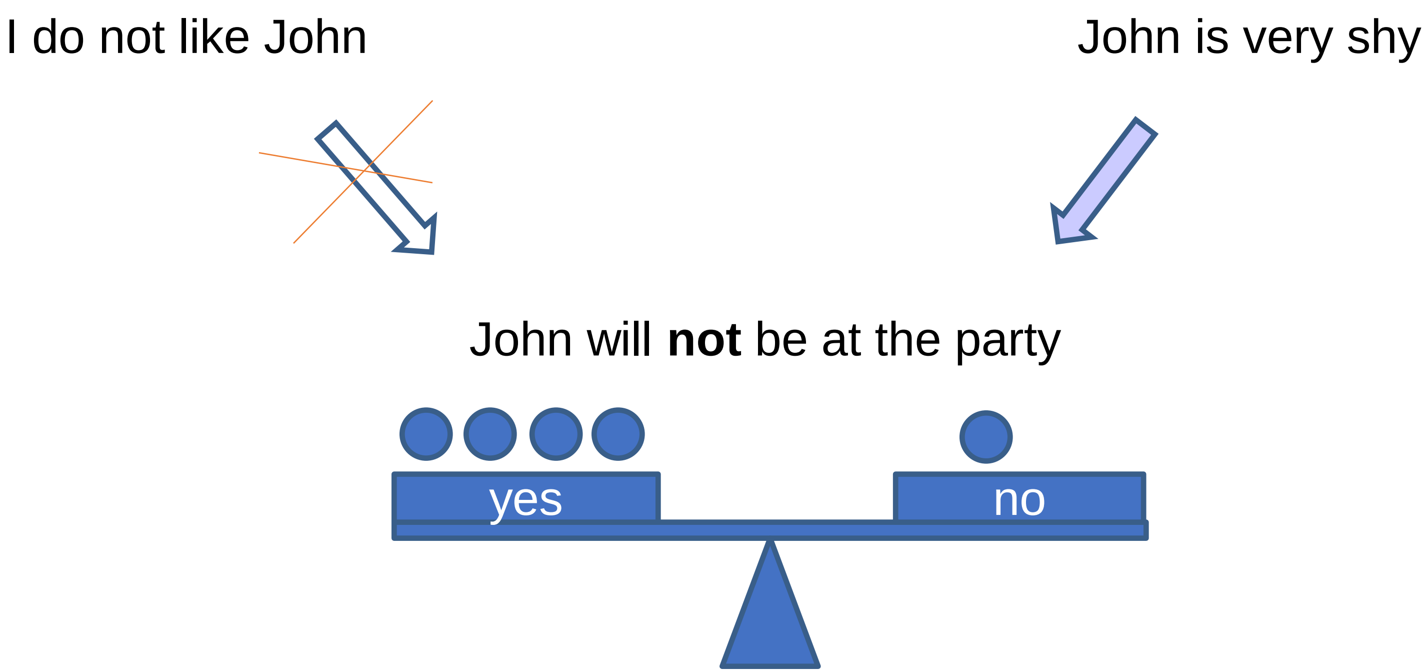
John will **not** be at the party



What are the odds?

I believe that John will not be at the party

Probabilistic reasoning: valid fact (evidence)



What are the odds given this fact?

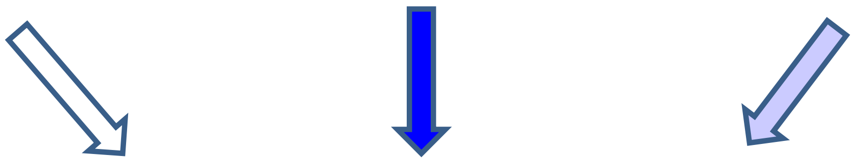
I believe that John will not be at the party

More facts – update your beliefs

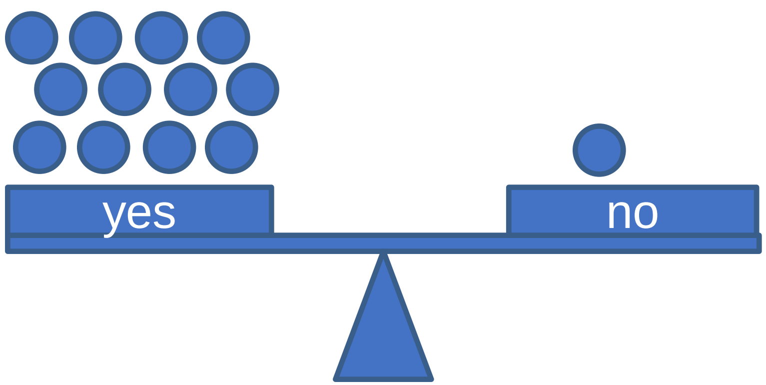
I do not like John

John is in Beijing

John is very shy



John will **not** be at the party



What are the odds?

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 evidence becomes available

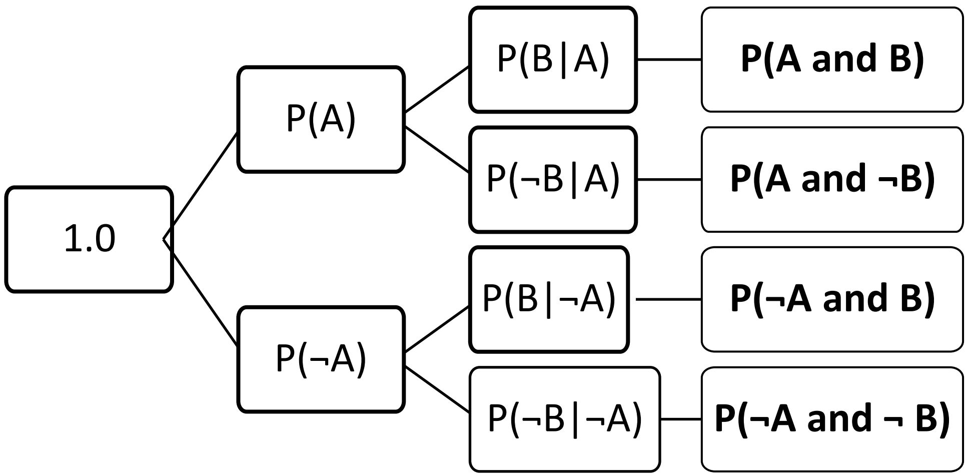
T. Bayes.
1701 - 1761

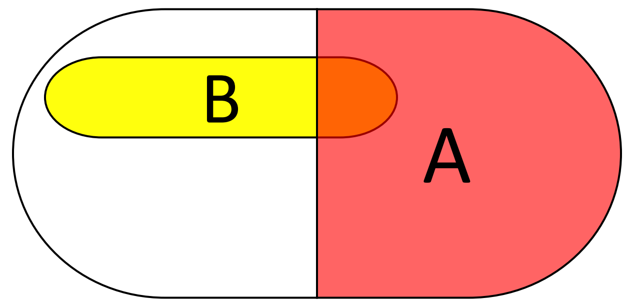
Bayes' method for updating beliefs

- There are 2 mutually exclusive 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.

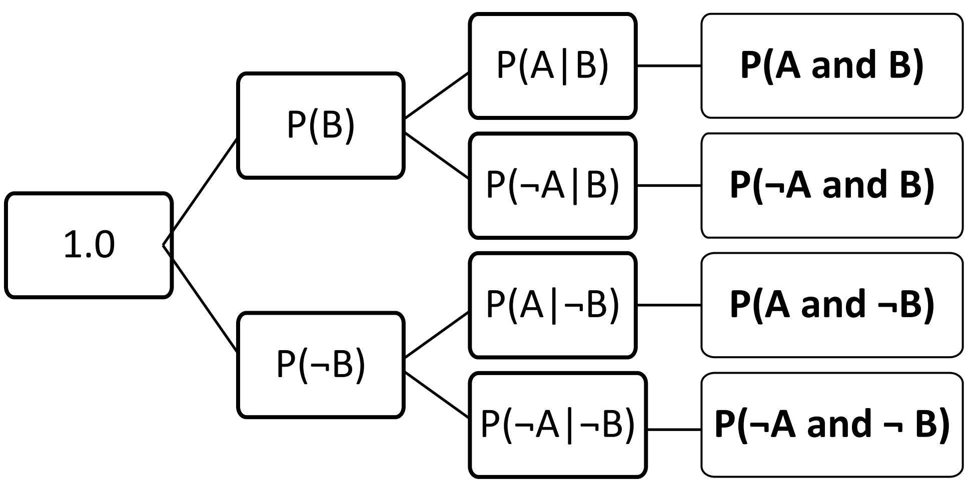
P(A) = 1/2, P(B) = 1/6

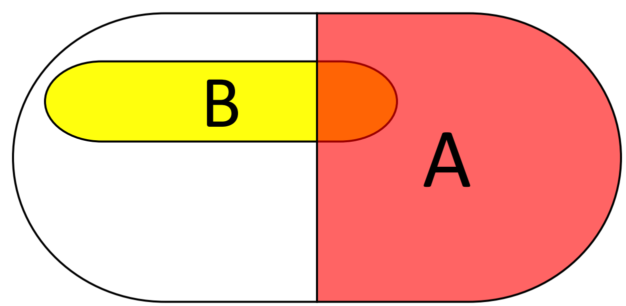
Two random events (not independent) happen at the same time – P(A and B)





Possible event combinations when we know the outcome of event A:
P(A)=1/2, P(B|A)=1/12 and P(A and B)=1/2*1/12 = 1/24

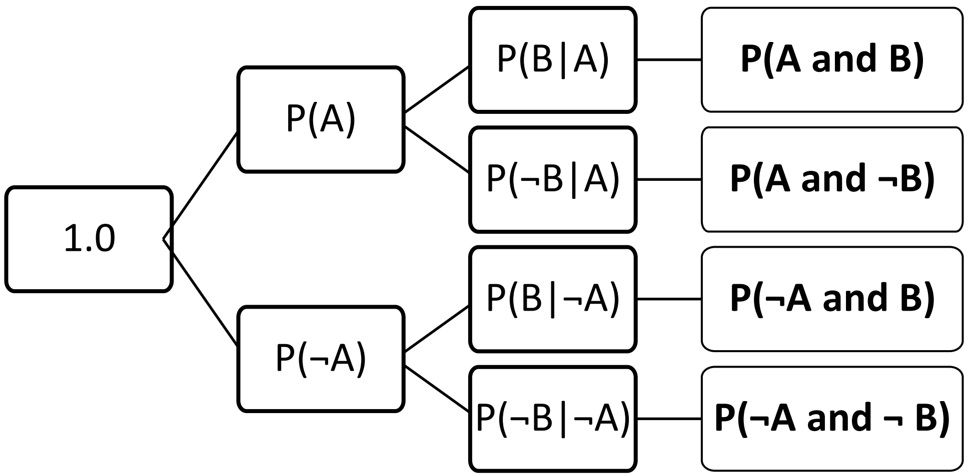




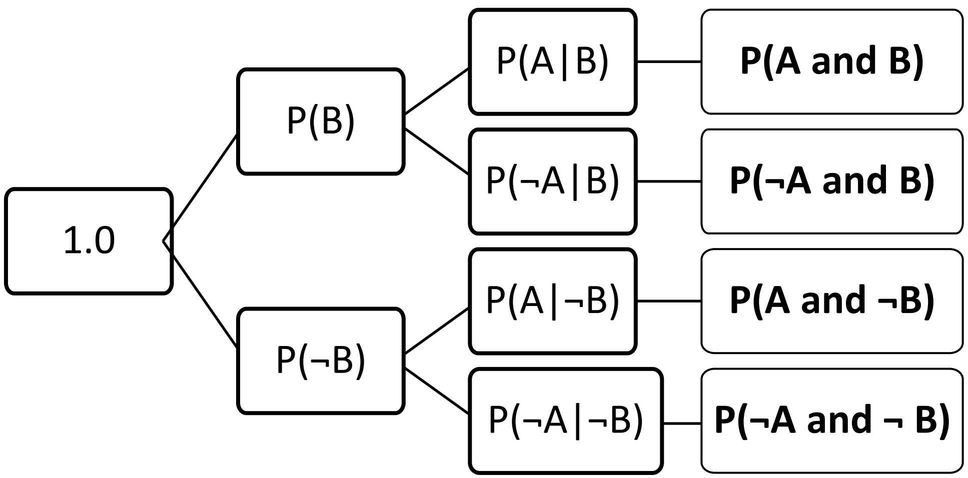
Possible event combinations when we know the outcome of event B:
P(B)=1/6, P(A|B)=1/4 and P(B and A)=1/6*1/4 = 1/24

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

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)

Bayes theorem

Bayes theorem (formalized by Laplace)

$$P(A|E) = P(A \cap E)/ P(E)$$
$$P(E|A) = P(A \cap E)/ P(A)$$

⇓

Probability of event A given evidence

Probability of evidence given event A

$$P(A|E) = P(E|A)P(A)/P(E)$$

Probability of event A without evidence (***prior probability***)

Inverse probabilities are typically easier to ascertain

Bayes method with probabilities

- 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.

The updated odds are computed as:

$$\frac{P(A|E)}{P(B|E)} = \frac{P(E|A)P(A)/P(E)}{P(E|B)P(B)/P(E)}$$

Bayes' method with probabilities

- 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.

or simply

$$\frac{P(A|E)}{P(B|E)} = \frac{P(E|A)P(A)}{P(E|B)P(B)}$$

Example: hit-and-run (fictitious)

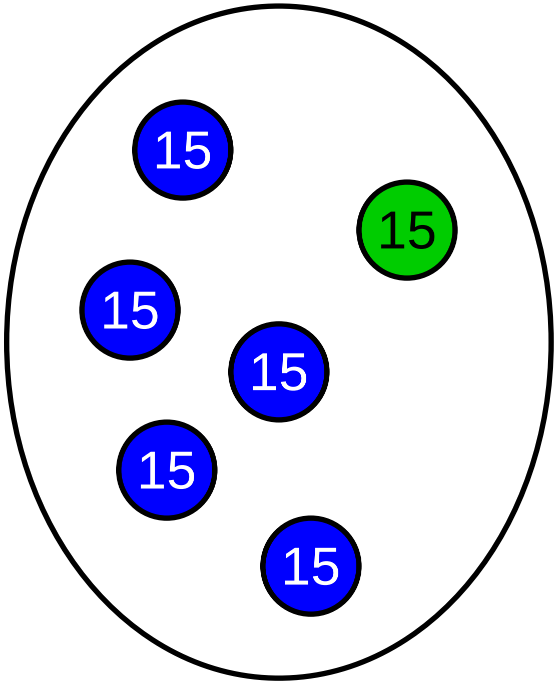
- Taxicab company has 75 blue cabs (**B**) and 15 green cabs (**G**)

- At night when there are no other cars on the street: hit-and-run episode

- Question: what is more probable:

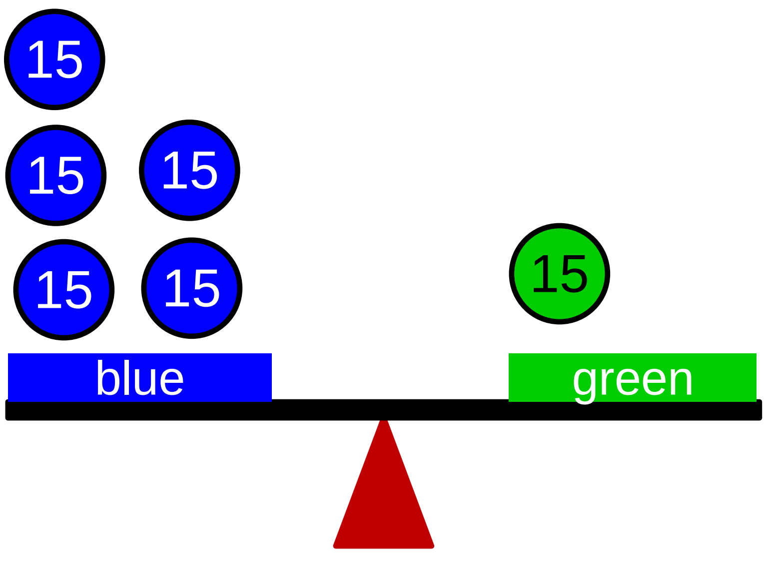
  B or **G**

  ?



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

What is more probable:
B or G



P(**B**):P(**G**)=5:1

New evidence

- Witness: “I saw a green cab”: E_G
- What is the probability that the witness really saw a green car?
- Witness is tested at night conditions: identifies correct color 4 times out of 5

- The eyewitness test shows:

$P(E_G \mid G) = 4/5$ (correctly identified)

$P(E_G \mid B) = 1/5$ (incorrectly identified)

Updating the odds

- In our case we want to compare:

the car was **G** given a witness testimony E_G: $P(G|E_G)$

vs.

the car was **B** given a witness testimony E_G: $P(B|E_G)$

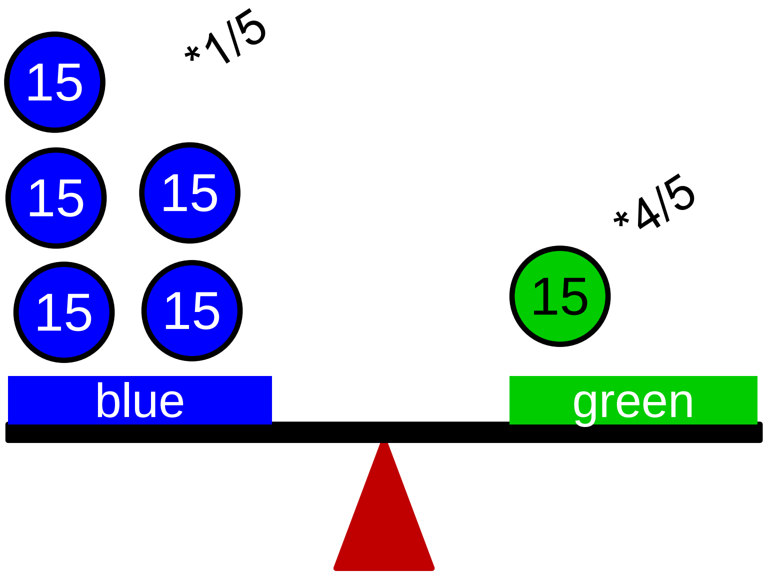
Note: There is no way to know which of 2 was true, we just estimate

Back to hit-and-run

All cabs were on the streets:

Prior odds ratio: P(**B**) : P(**G**) = 5/1

Updated odds ratio: $\frac{P(B|E_G)}{P(G|E_G)} = \frac{P(B)*P(E_G|B)}{P(G)*P(E_G|G)}$



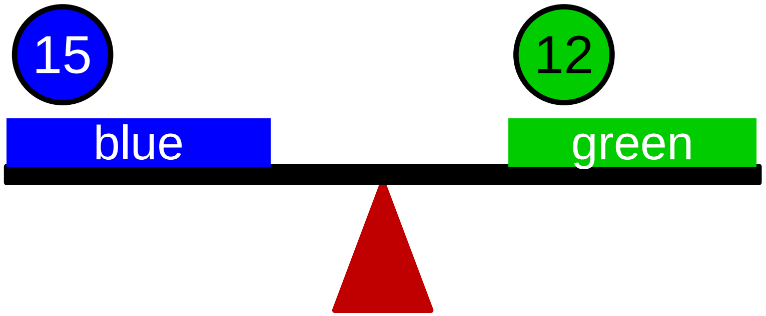
P(E_G|**G**)= 4/5 (correctly identified)

P(E_G|**B**)= 1/5 (incorrectly identified)

New odds

$$\frac{P(B|E_G)}{P(G|E_G)} = \frac{P(B)*P(E_G|B)}{P(G)*P(E_G|G)}$$

Still 5:4 odds that the car was **B**!



Hit-and-run: full calculation

$P(\mathbf{B}) = 5/6$, $P(\mathbf{G}) = 1/6$

$P(\mathbf{E_G} \mid \mathbf{G}) = 4/5$ $P(\mathbf{E_G} \mid \mathbf{B}) = 1/5$

- Probability that car was **green** given the evidence E_G:

$P(\mathbf{G}|\mathbf{E_G}) = P(\mathbf{G}) * P(\mathbf{E_G}|\mathbf{G}) / P(\mathbf{E_G}) = [1/6 * 4/5] / P(E_G) \qquad = 4/30P(E_G)$

//- 4 parts of $30P(X_G)$

- Probability that car was **blue** given the evidence X_G:

$P(\mathbf{B}|\mathbf{E_G}) = P(\mathbf{B}) * P(\mathbf{E_G}|\mathbf{B}) / P(\mathbf{E_G}) = [5/6 * 1/5] / P(\mathbf{E_G}) = 5/30P(E_G)$

//- 5 parts of $30P(X_G)$

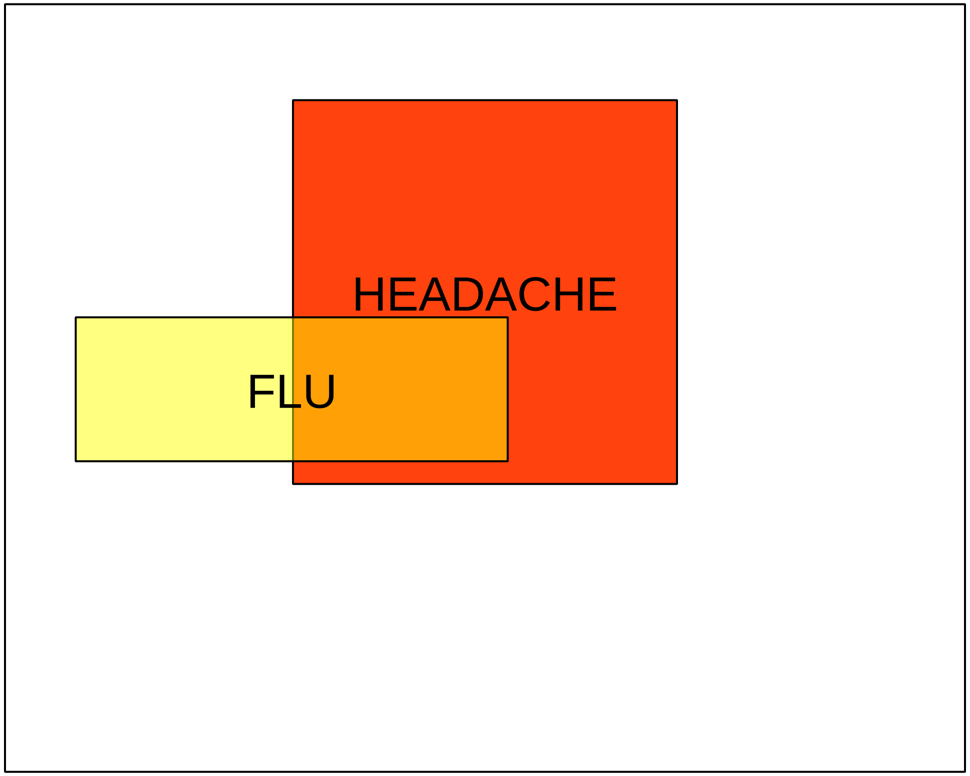
Bayes in ‘real’ life. Example

P(H)=1/10

P(F)=1/40

P(H|F)=1/2

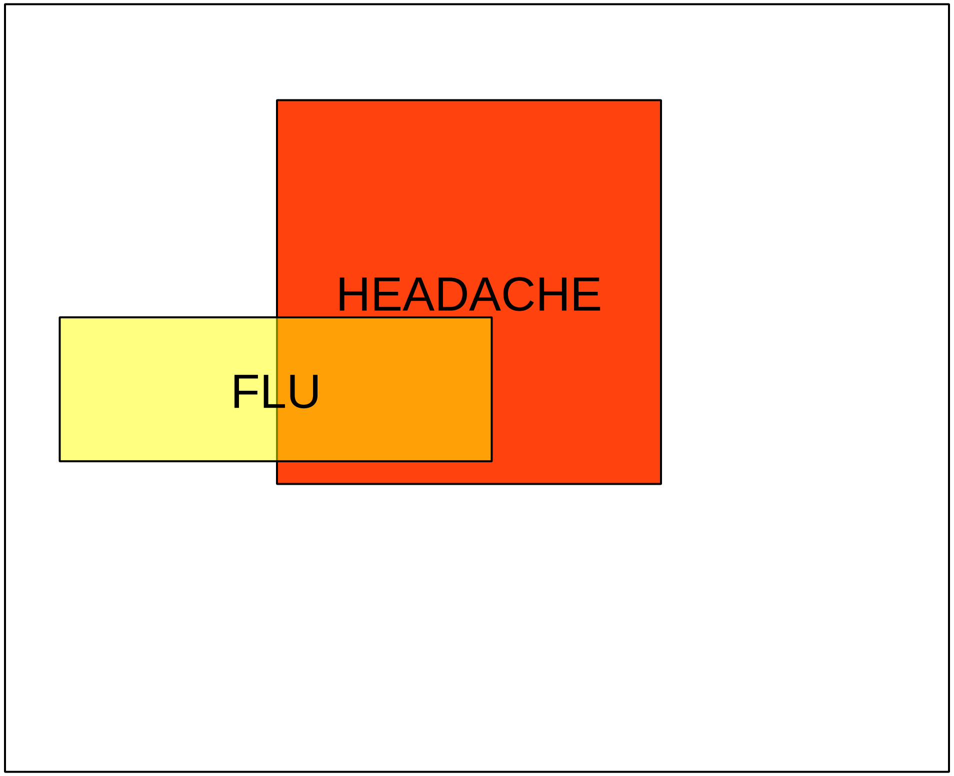
P(F|H) =?



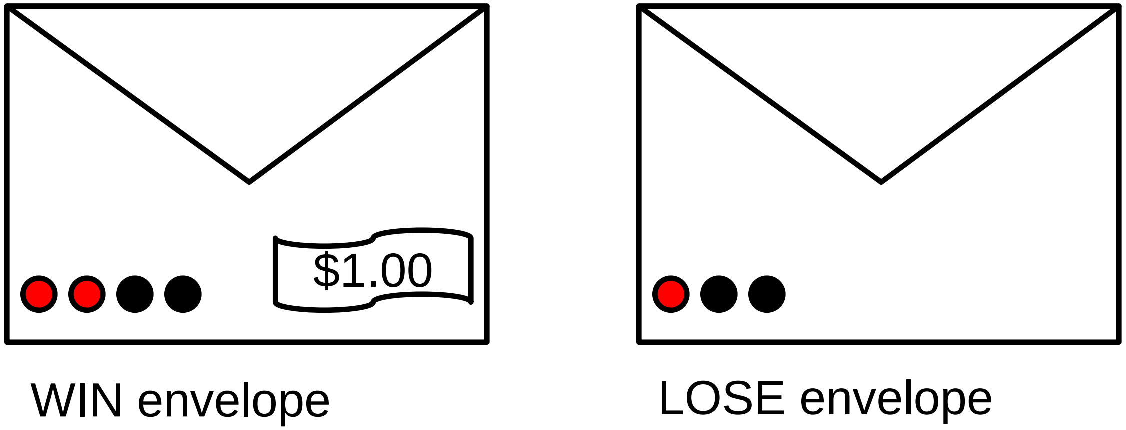
Bayes in 'real' life. Example

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

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



Bayes in 'real' life. Activity

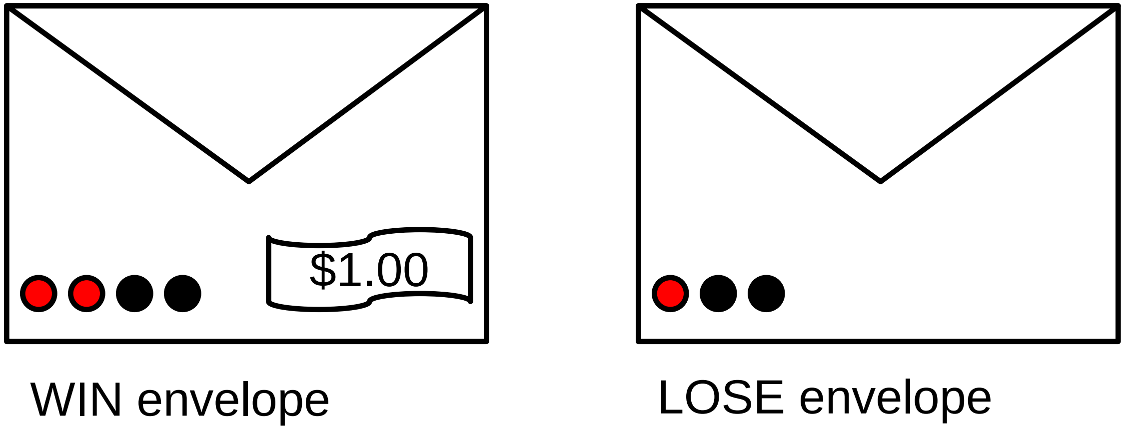


Someone draws an envelope at random and offers to sell it to you.

How much should you pay?

The probability to win is 1:1. Pay no more than 50c.

Bayes in 'real' life. Activity



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?

When you want to:

- Determine the probability of having a medical condition after positive test results
- Find out a probable outcome of political elections
- Improve machine-learning performance
- Even to “prove” or “disprove” the existence of God

Use Bayesian Reasoning

Summary: 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 (¬A|B), i.e. given evidence B what probability is higher: that A occurred or that ¬A occurred?

- We know P(A) and P(¬A) – prior probabilities
- We know P(B|A) and P(B|¬A)

- From Bayes' theorem:

$$\mathbf{P(A|B)} = P(A)*P(B|A) / P(B)$$

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

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

$$\frac{P(A|B)=P(B|A)P(A)/P(B)}{P(\neg A|B)=P(B|\neg A)P(\neg A)/P(B)}$$

- Instead of finding ratio, to avoid underflow, use log:

$$\log \frac{P(A|B)=P(B|A)P(A)}{P(\neg A|B)=P(B|\neg A)P(\neg A)}$$

If positive then A is more probable, if negative then ¬A is more probable