Conditional Probability
Bayes Theorem

Lecture 6

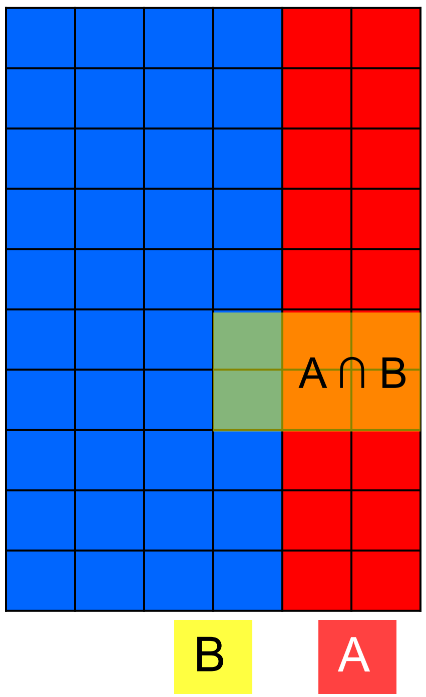
Conditional probability

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

$\#(A \cap B) = 4$

Out of $\#(B) = 6$

$P(A|B) = 4/6 = 2/3$



- Formally: $P(A|B)= P(A \cap B) / P(B)$

$P(A) = 20/60=1/3$

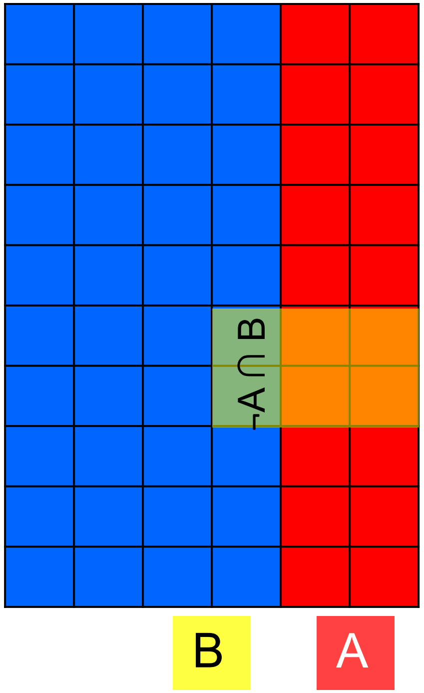
$P(B) = 6/60$

$P(A \cap B) = 4/60$

$P(A|B) = (4/60) / (6/60) = 2/3$

Conditional probability

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



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

$P(\neg A|B) = (2/60)\ /\ (6/60) = 2/6$

Probabilistic independence: formal definition

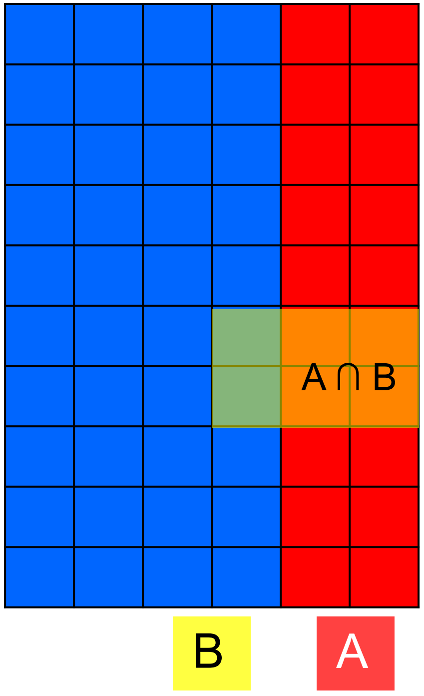
- Two events A and B are *mutually independent* if $P(A|B) = P(A)$, which means that:

  $P(A|B)$ is the same as $P(A)$

- In other words, knowing that B is true (or false) does not change the probability of A

Example 1: A and B are not independent

- $P(B) \neq P(B|A)$
- Knowing that A is true (A has occurred) changes the probability of B



$P(B) = 6/60 = \mathbf{0.1}$ unconditional

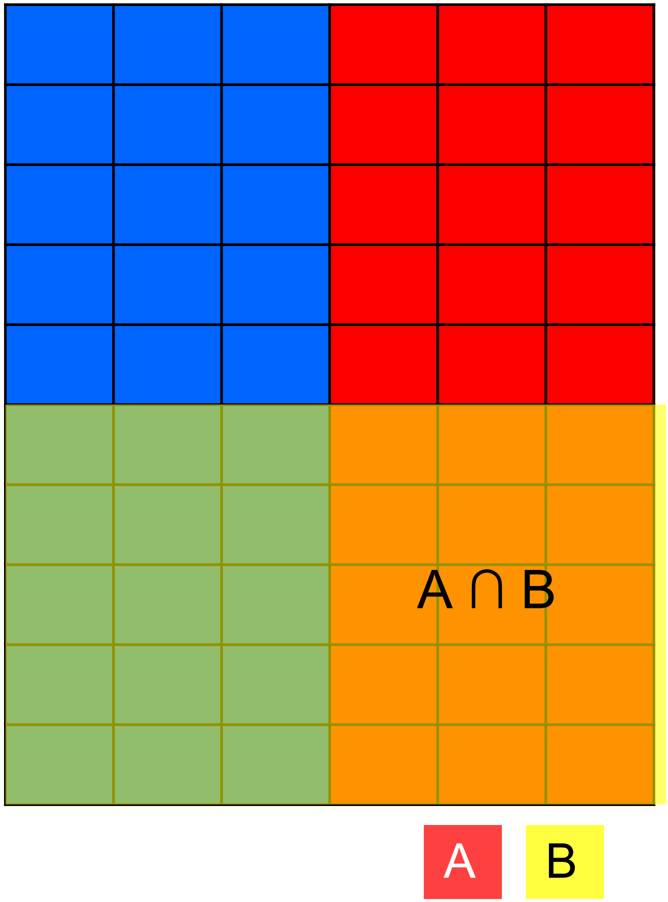
$P(B|A) = (4/60) / (20/60) = 4/20 = 1/5 = \mathbf{0.2}$ conditional

Example 2: A and B are independent

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

$P(A|\neg B) = P(A)$ $15/30 = 30/60$

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

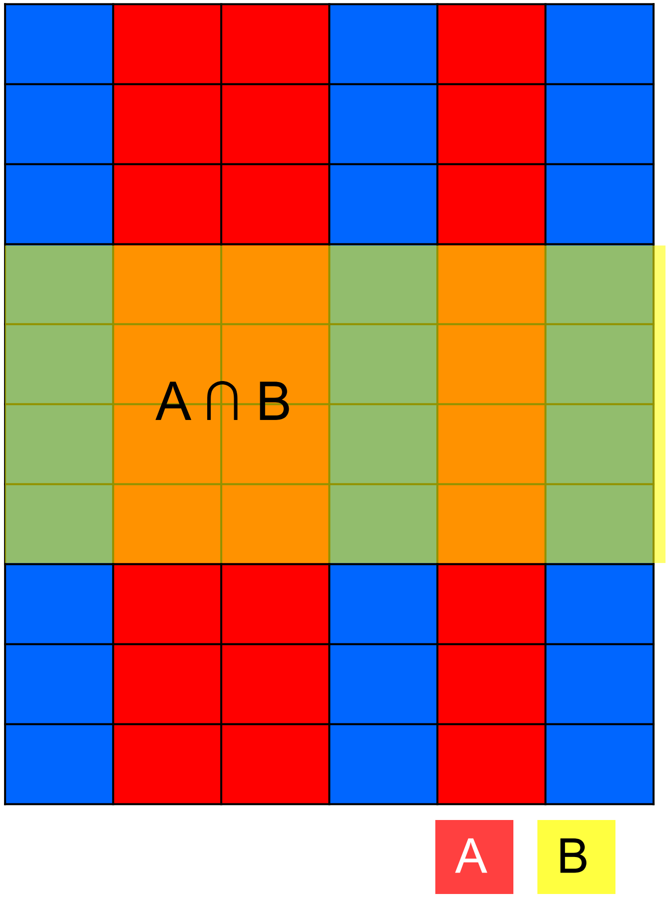


Example 3: A and B are independent

P(A|B) = P(A) 12/24 = 30/60

P(A| ¬ B) = P(A) 18/36 = 30/60

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



Probabilistic independence is not the same as a colloquial independence!

- Colloquial independence: “An *independent* committee reevaluated the test results” - no independence in statistical sense
- Statistical independence: needs a check whether $P(A|B) \overset{?}{=} P(A)$

Example 3: rolling two dice

Consider 2 dice events:

- A = ‘the event of getting 6 on the first throw’.
- B = ‘the event of getting 6 on the second throw’.

Are A and B independent?

Example 3: rolling two dice

Consider 2 dice events:

- A = ‘the event of getting 6 on the first throw’.
- B = ‘the event of getting 6 on the second throw’.

Are A and B independent?

$$\left.\begin{aligned} P(B) &= 1/6 \\ P(B|A) &= 1/6 \end{aligned}\right\} \Rightarrow \text{A and B are independent}$$

Example 4: rolling two dice

Consider 2 dice events:

- A = “the event of getting 6 on the first throw”.
- B = “the event of getting the sum of 8 on two throws”.

Are A and B independent?

Example 4: rolling two dice

Consider 2 dice events:

- A = ‘the event of getting 6 on the first throw’.
- B = ‘the event of getting the sum of 8 on two throws’.

Are A and B independent?

P(B) = 5/36

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

P(B/A) = 1/6

	6
1	7
2	8
3	9
4	10
5	11
6	12

=> A and B are **NOT** independent

Now you do it:
are A and B independent?

Consider 2 dice events:

- A = ‘number on a dice is a multiple of 2’.
- B = ‘number on a dice is a multiple of 3’.

Now you do it:
are A and B independent?

Consider 2 dice events:

- A = 'number on a dice is a multiple of 2'.
- B = 'number on a dice is a multiple of 3'.

P(B) = 2/6

P(A) = 3/6

P(B|A) = 1/3

given that the number is already a multiple of 2, what is the probability that it is also a multiple of 3

P(A|B) = 1/2

<u>P(B|A) = P(B)</u>

<u>P(A|B) = P(A)</u>

=> A and B are independent

Dependent events imply correlation (which is **not causation**)

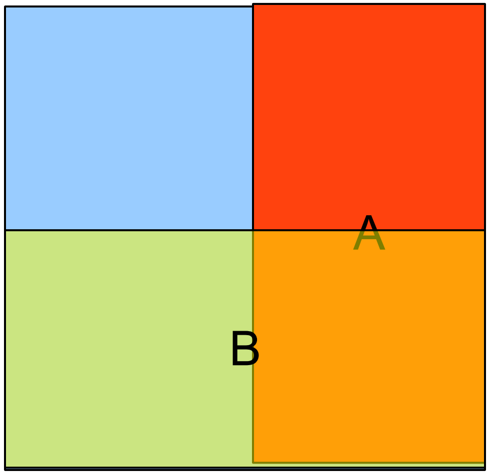
- If $P(A|B) > P(A)$ (or $P(A|B) < P(A)$), then A and B are not independent
- There is a relationship between A and B: if B occurred, then A becomes more (or less) probable (conditional **dependence**)
- This relationship implies *correlation* between A and B (but does not imply causation)
- **Example**: P[Disease | Visiting a doctor] > P[Disease]

  This does not imply that visiting a doctor increases the probability of the disease!

Independent vs. mutually exclusive

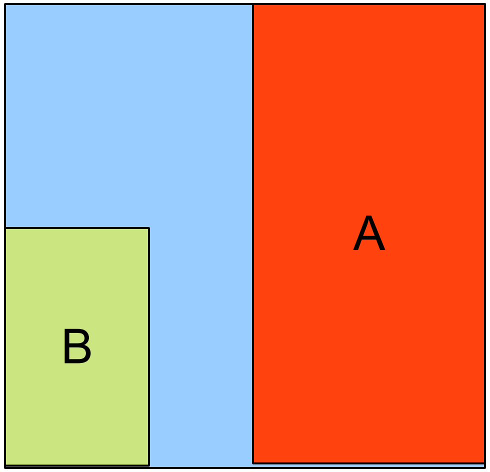
A is ***independent*** of B: knowing that B is true (or false) does not change the probability of A:

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



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

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



Joint probability of two events

From the definition of conditional probabilities:

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

we can compute $P(A \cap B)$ – the **joint probability** of any two events A and B (intersection, both events happened together):

$$\mathbf{P(A \cap B)} = P(A|B)\, P(B)$$

Joint probability of independent events

The **joint probability** of A and B:

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

If A and B are *independent* that becomes:

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

Theorem of Probability 3. The multiplication rule
(Intersection of two independent events)

Joint probability of mutually exclusive events

If A and B are mutually exclusive, then their intersection is empty:

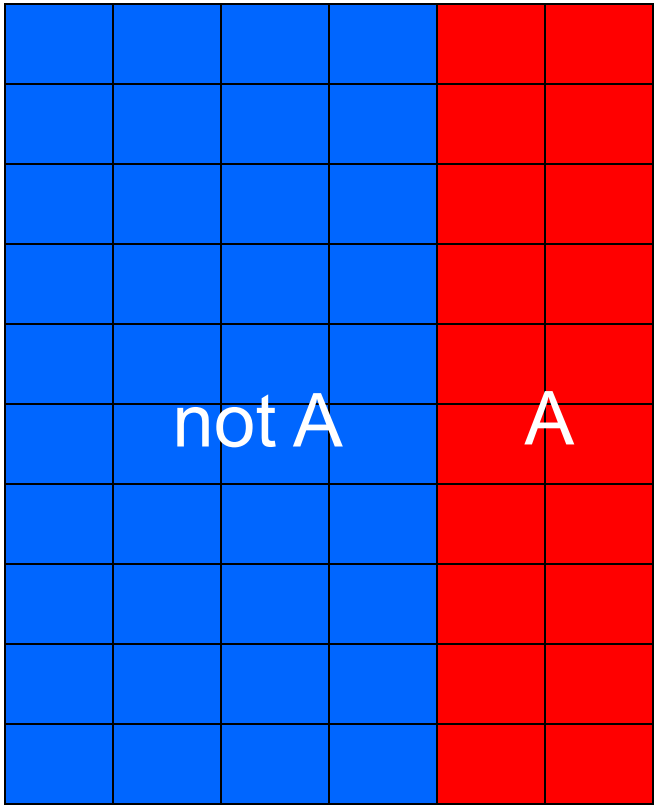
$P(A \cap B) = 0$

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

Unions:

$P(A \cup B) = P(A)+P(B)$

$P(A \cup \neg A) = P(A) + P(\neg A) = 1$



Bayesian reasoning

Updating beliefs with evidence

Inductive reasoning with probabilities

- Develop critical thinking: always have good **reasons** for your beliefs
- "Some reasons are 100% correct" – never 100%!

- Two types of logical reasoning:
  - *Deductive*: logical process where conclusions are drawn from general principles or premises
  - *Inductive*: building general conclusions from facts (data)

- In statistics we mostly work with inductive reasoning
- Inductive conclusions are not guaranteed to be true
- They work with probabilities => we always have a chance of being wrong!

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 the ***odds*** of A vs. B.
- Collect evidence data **E.**
- Re-estimate P(**A**|**E**):P(**B**|**E**) and update your beliefs.

Basis of Bayesian reasoning

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

From definition of Conditional probability:
$P(A|B) = P(A \cap B) / P(B)$

On the other hand:

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

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

Now we can express conditional probability of A given B through conditional probability of B given A and unconditional probabilities of A and B:

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

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

$$\boxed{P(A|E) = P(E|A) / P(E) * P(A)}$$

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

Inverse probabilities are typically easier to ascertain

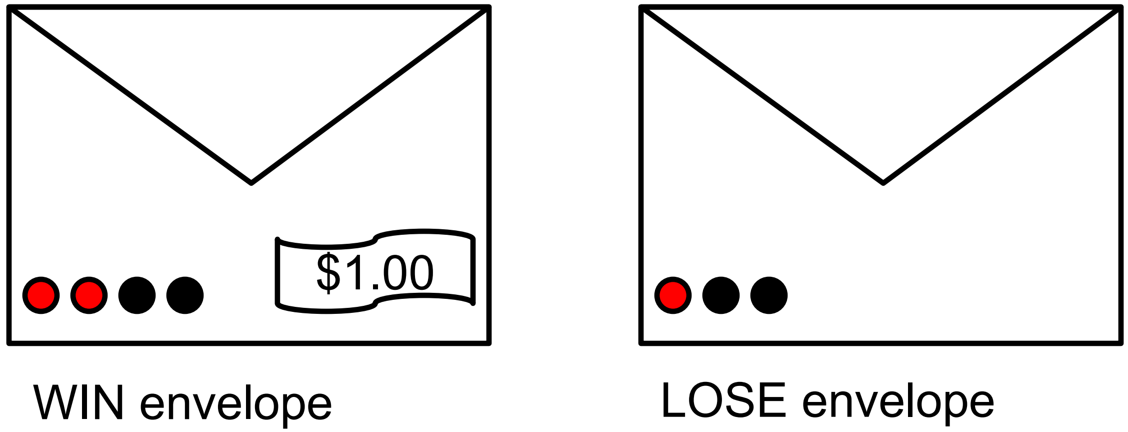
Updating odds given evidence

- 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 prior 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:

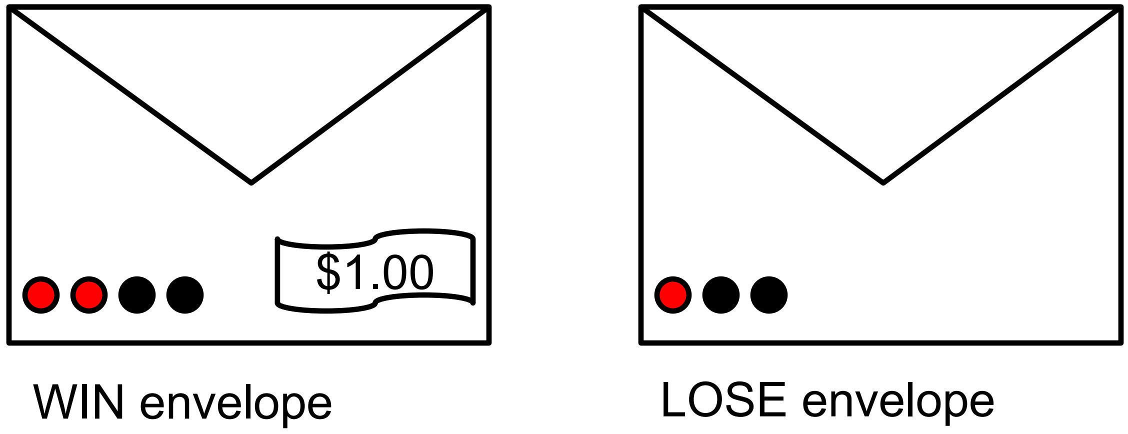
$$\frac{P(A|E)}{P(B|E)} = \frac{P(E|A)/P(E) * P(A)}{P(E|B)/P(E) * P(B)} \quad => \quad \frac{P(A|E)}{P(B|E)} = \frac{P(E|A)P(A)}{P(E|B)P(B)}$$

Example 5: envelopes



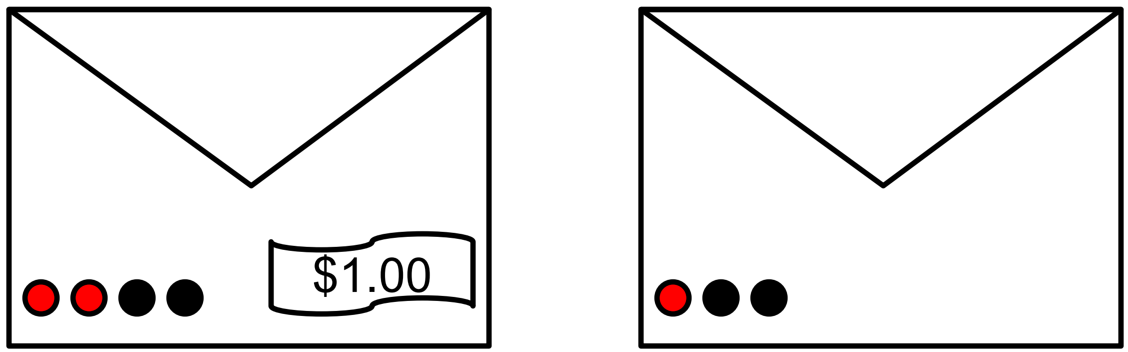
Someone draws an envelope at random and offers to sell it to you.
How much should you pay?

Example 5: envelopes



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.

Example 5: envelopes

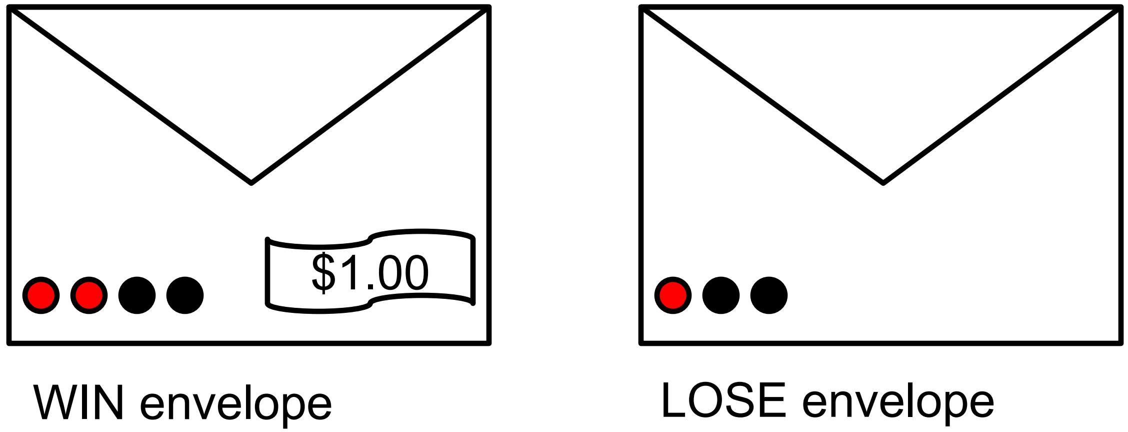


WIN envelope LOSE envelope

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

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

Example 5: envelopes



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

Example 6:
hit-and-run (fictitious)

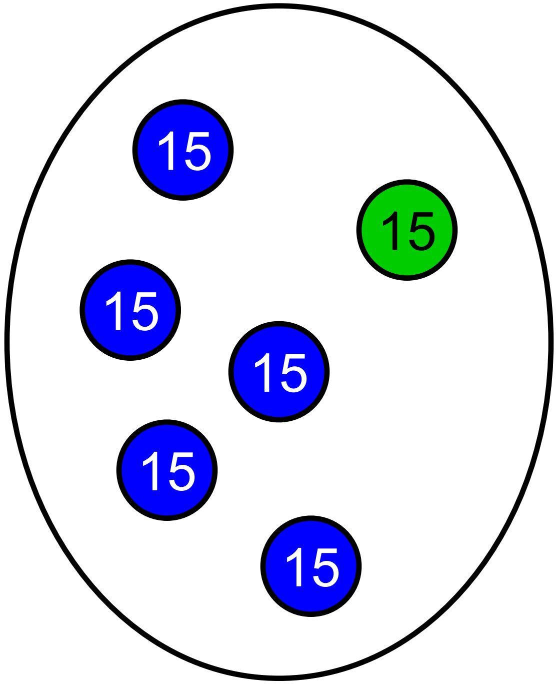
- Taxicab company has 75 blue cabs (**B**) and 15 green cabs (**G**)

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

- Question: what is more probable:

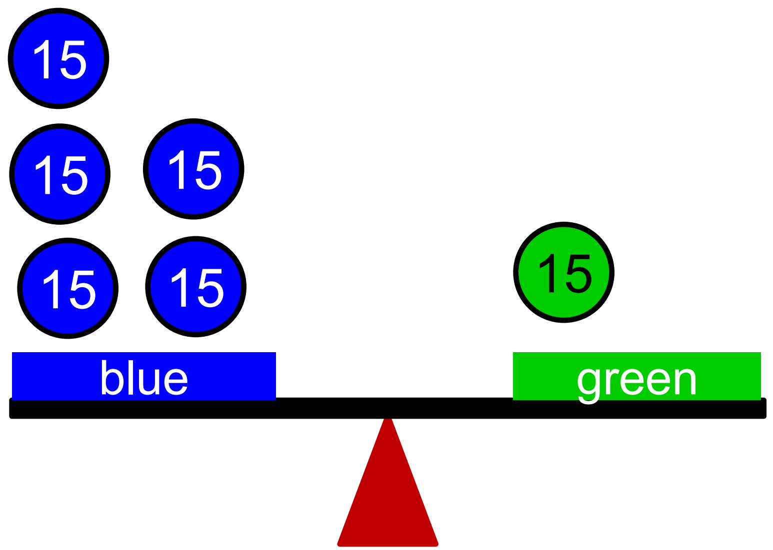
  B or **G**

  ?



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

In absence of evidence – 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 and 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)

=> **a pretty reliable witness**

Updating the odds

- In our case we want to compare*:

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

vs.

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

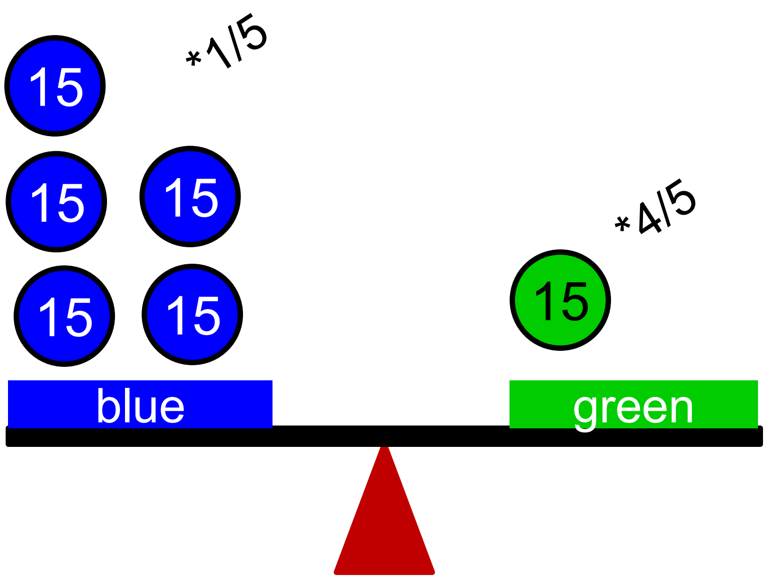
*Disclaimer: There is no way to know which of 2 was true, we just *estimate*

Back to hit-and-run

All cabs were on the streets:

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

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



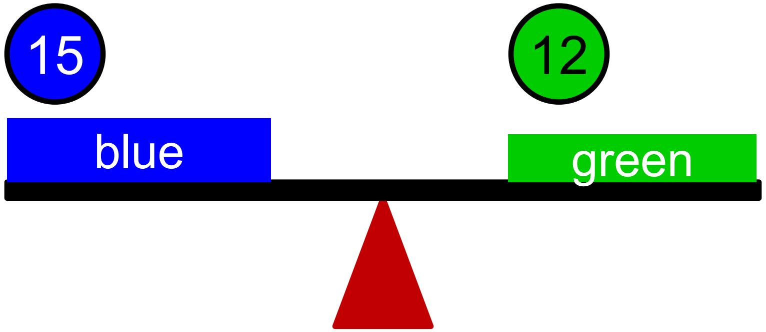
P($\mathbf{E_G}$ | **G**) = 4/5 (correctly identified)

P($\mathbf{E_G}$ | **B**) = 1/5 (incorrectly identified)

Updated odds

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

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



Hit-and-run: full calculation

$P(B) = 5/6$, $P(G) = 1/6$

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

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

$P(G|E_G) = P(G) * P(E_G|G) / P(E_G) = [1/6 * 4/5] / P(E_G) \quad = 4/30P(E_G)$

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

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

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

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

Exercise: probability of rain

- You're about to get on a plane to Seattle for a one-day interview at Amazon. You want to know if you should bring an umbrella.
- The probability that it's raining on any given day in Seattle is 25%.
- You call a friend of yours who lives there and ask him if it's raining. He tells you that "Yes, it is raining".

Based on history, your friend has a 2/3 chance of telling you the truth and a 1/3 chance of messing with you by lying.

- What is the probability that it is actually raining in Seattle?

Exercise: probability of rain

- You're about to get on a plane to Seattle for a one-day interview at Amazon. You want to know if you should bring an umbrella. The probability that it's raining on any given day in Seattle is 25%.

Prior probabilities: P(Rain) = 0.25, P(¬ Rain) = 0.75

- You call a friend of yours who lives there and ask him if it's raining. He tells you that "Yes, it is raining”. Based on history, your friend has a 2/3 chance of telling you the truth and a 1/3 chance of messing with you by lying.
- We have evidence E_{Rain} from one friend (call it E)
- Based on the friend: $P(E_{Rain} \mid Rain) = 2/3$ (telling the truth), $P(E_{Rain} \mid \neg Rain) = 1/3$
- What is the probability that it is actually raining in Seattle? Let’s rename $E_{Rain} \rightarrow E$

P(Rain | E) = P(E | Rain)*P(Rain)/P(E) = 2/3 * 0.25/P(E)

P(¬Rain | E) = P(E | ¬Rain)*P(¬Rain)/P(E) = 1/3 * 0.75/P(E)

Odds P(Rain | E) : P(¬Rain| E) = (2/3*1/4) / (1/3*3/4) = 2/12*12/3`=2:3

P(Rain | E) + P(¬Rain | E) = 1

P(Rain | E) = 2/5 = 0.4 (despite the fact that your friend tells you the opposite)

Bayes: common applications

- A common application of Bayes' rule is in diagnostic and testing.
- Possible application: legal reasoning
  - That is debated <= people are not well educated in Bayesian reasoning
  - *Prosecutor's Fallacy*: a common error where the probability of evidence given the defendant is guilty is confused with the probability of the defendant being guilty given the evidence
- In Machine Learning: classification

Example 7: diagnosing meningitis

- A doctor knows that 50% of patients which have meningitis have a symptom of a stiff neck.
- Stiff neck is not a rare symptom and occurs in a general population with probability 1/20

- The **doctor also knows some unconditional** facts (prior probabilities):

  the prior probability that any patient **has** meningitis is 1/50,000

  => the probability that he **does not have** a meningitis is 49,999/50,000

- A patient comes to the office with a stiff neck.
- What is the probability that he has a meningitis?

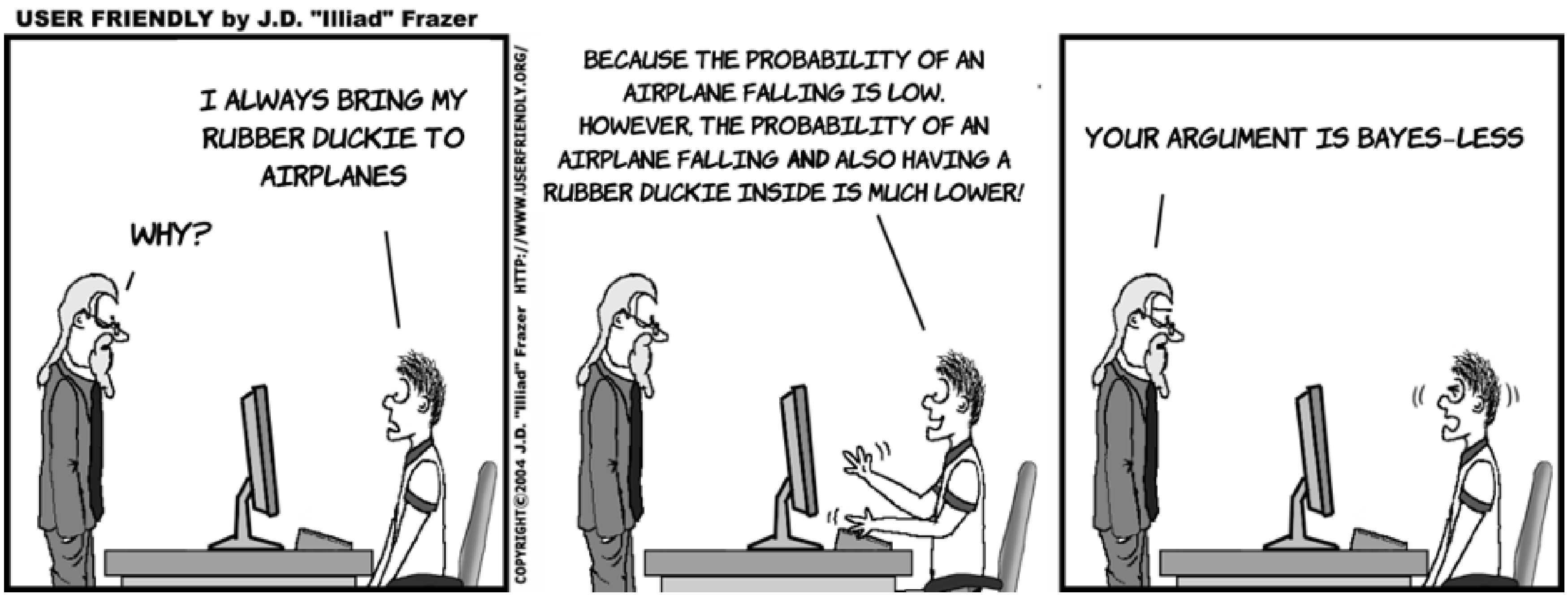
Probability of having Meningitis

P(StiffNeck| Meningitis) = 0.5
P(StiffNeck) = 1/20 = 0.05
P(Meningitis) = 1/50000

P(Meningitis | StiffNeck)
= P(StiffNeck | Meningitis) P(Meningitis) / P(StiffNeck)
= (0.5) x (1/50000) / 0.05 =0.0002

Bayes' rule critics: prior probabilities

- The doctor had the following quantitative information in the diagnostic direction from symptoms (evidences, effects) to causes:
  P(StiffNeck| Meningitis) = 0.5, P(Meningitis) = 1/50000
- The problem is that prior probabilities are hard to estimate, and they may fluctuate.
- Imagine, there is a sudden epidemic of meningitis. The prior probability, P(Meningitis), will go up.
- Clearly, P(StiffNeck|Meningitis) is unaffected by the epidemic. It simply reflects the way meningitis works.
- The estimation of P(Meningitis|StiffNeck) will remain incorrect until new data about P(Meningitis) are collected


USER FRIENDLY by J.D. "Illiad" Frazer
I ALWAYS BRING MY RUBBER DUCKIE TO AIRPLANES
WHY?
COPYRIGHT©2004 J.D. "Illiad" Frazer HTTP://WWW.USERFRIENDLY.ORG/
BECAUSE THE PROBABILITY OF AN AIRPLANE FALLING IS LOW. HOWEVER, THE PROBABILITY OF AN AIRPLANE FALLING AND ALSO HAVING A RUBBER DUCKIE INSIDE IS MUCH LOWER!
YOUR ARGUMENT IS BAYES-LESS

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