Introduction to the likelihood ratio framework for evaluation of forensic evidence

Geoffrey Stewart Morrison

Forensic Data Science Laboratory
Aston University

Aston University
BIRMINGHAM UK

Workshop material

- Slides

https://forensic-data-science.net/#EAFS2025

Additional training

- Concepts of forensic inference and statistics
    - Master's level continuing professional development course
    - Online delivery with weekly interactive sessions
    - Delivered in 22 weeks spread over 6 months
    - ~1 day per week workload
    - Competency assessment

https://www.aston.ac.uk/study/courses/concepts-forensic-inference-and-statistics-standalone-module/

Content

- Bayesian Reasoning
- Similarity and Typicality
- Conditional Probabilities
- Bayes' Theorem Part I: Prior odds, likelihood ratios, & posterior odds
- Bayes' Theorem Part II: Responsibilities
- Bayes' Theorem Part III: Updating beliefs
- Illogical Reasoning

Bayesian Reasoning

Imagine you are driving to the airport...

Imagine you are driving to the airport...

Imagine you are driving to the airport...

Imagine you are driving to the airport...

initial probabilistic belief & evidence → updated probabilistic belief

higher?
or
lower?

Imagine you are driving to the airport...

initial probabilistic belief	&	evidence	→	updated probabilistic belief

higher?
or
lower?

Bayesian reasoning

- What we have just been doing is Bayesian reasoning
- It is about logic
- It is not about mathematical formulae or databases
- There is nothing complicated or unnatural about it
- It is the logically correct way to think about many problems

Thomas Bayes?

Pierre-Simon Laplace

Similarity and Typicality

Imagine you work at a shoe recycling depot ...

- You pick up two shoes of the same size
    - Does the fact that they are of the same size mean they were worn by the same person?

Imagine you work at a shoe recycling depot ...

- You pick up two shoes of the same size
    - Does the fact that they are of the same size mean they were worn by the same person?
    - Does the fact that they are of the same size mean that it is highly probable that they were worn by the same person?

Imagine you work at a shoe recycling depot ...

- You pick up two shoes of the same size
  - Does the fact that they are of the same size mean they were worn by the same person?
  - Does the fact that they are of the same size mean that it is highly probable that they were worn by the same person?

Both
similarity
and
typicality
matter

Imagine you work in a forensic footwear-comparison laboratory ...

suspect's
shoe

crime-scene
shoemark

Imagine you work in a forensic footwear-comparison laboratory ...

- The shoemark at the crime scene is size 10

- The suspect’s shoe is size 10
  - What is the probability that the shoemark at the crime scene would be size 10 if it had been made by the suspect’s shoe?

Imagine you work in a forensic footwear-comparison laboratory ...

- The shoemark at the crime scene is size 10

- The suspect’s shoe is size 10
  - What is the probability that the shoemark at the crime scene would be size 10 if it had been made by the suspect’s shoe?

- Half the shoes at the recycling depot are size 10
  - What is the probability that the shoemark at the crime scene would be size 10 if it had been made by the someone else’s shoe?

Imagine you work in a forensic footwear-comparison laboratory ...

- The shoemark at the crime science and the suspect's shoe are both size 10

similarity / **typicality** = **1** / **0.5** = **2**

you are twice as likely to get a size 10 shoemark at the crime scene if it were produced by the suspect's shoe than if it were produced by someone else's shoe

- someone else selected at random from the relevant population

Imagine you work in a forensic footwear-comparison laboratory ...

- The shoemark at the crime scene is size 14

- The suspect's shoe is size 14
  - What is the probability that the shoemark at the crime scene would be size 14 if it had been made by the suspect's shoe?

Imagine you work in a forensic footwear-comparison laboratory ...

- The shoemark at the crime scene is size 14

- The suspect’s shoe is size 14
  - What is the probability of the shoemark at the crime scene would be size 14 if it had been made by the suspect’s shoe?

- 1% of the shoes at the recycling depot are size 14
  - What is the probability that the shoemark at the crime scene would be size 14 if it had been made by the someone else’s shoe?

Imagine you work in a forensic footwear-comparison laboratory ...

- The shoemark at the crime science and the suspect's shoe are both size 14

similarity / typicality = 1 / 0.01 = 100

you are 100 times more likely to get a size 14 shoemark at the crime scene if it were produced by the suspect's shoe than if it were produced by someone else's shoe

- someone else selected at random from the relevant population

Similarity and typicality

- **size 10**

  similarity / typicality = 1 / 0.5 = 2

- **size 14**

  similarity / typicality = 1 / 0.01 = 100

- If you didn't have a database, could you have made subjective estimates of relative proportions of different shoe sizes in the population and applied the same logic to arrive at a conceptually similar answer?

similarity / typicality = likelihood ratio

¿Area?

Logic

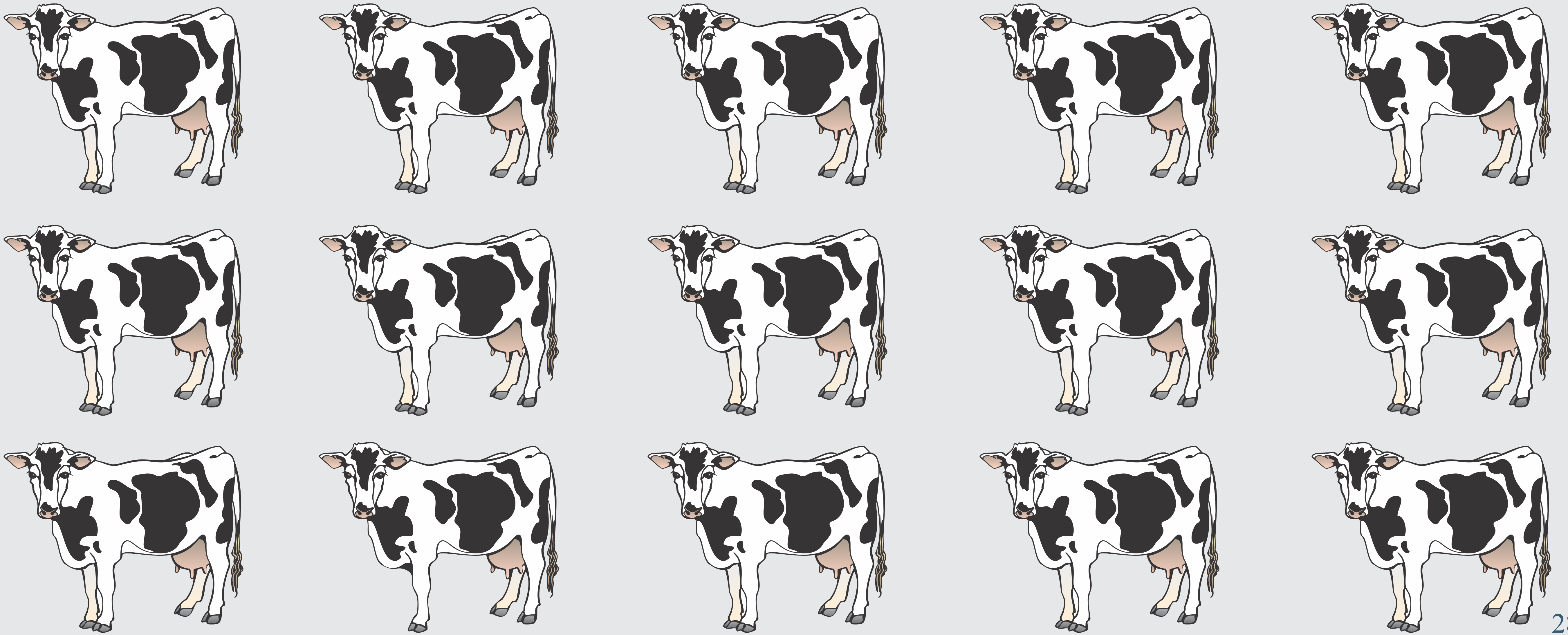
length × width = area

similarity / typicality = likelihood ratio

Conditional Probabilities

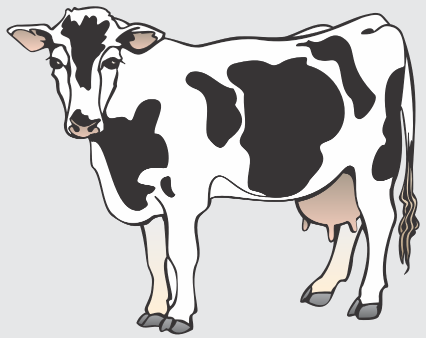
Given that it is a cow, what is the probability that it has four legs?

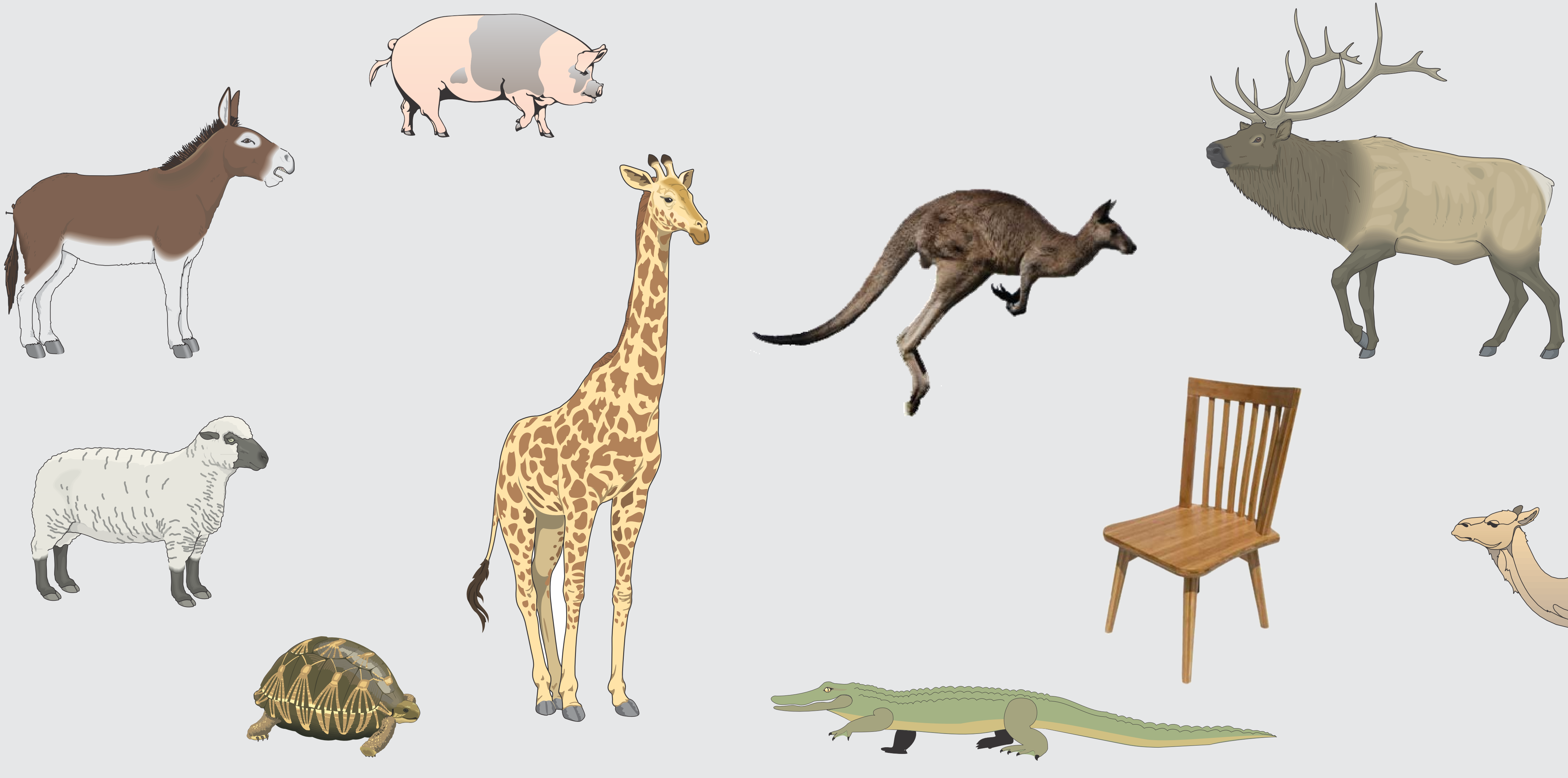
p(4 legs | cow) = ?

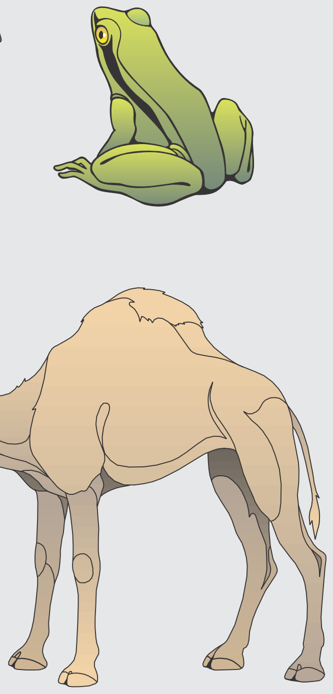


Given that it has four legs, what is the probability that it is a cow?

$$p(\, \text{cow} \mid \text{4 legs} \,) = ?$$







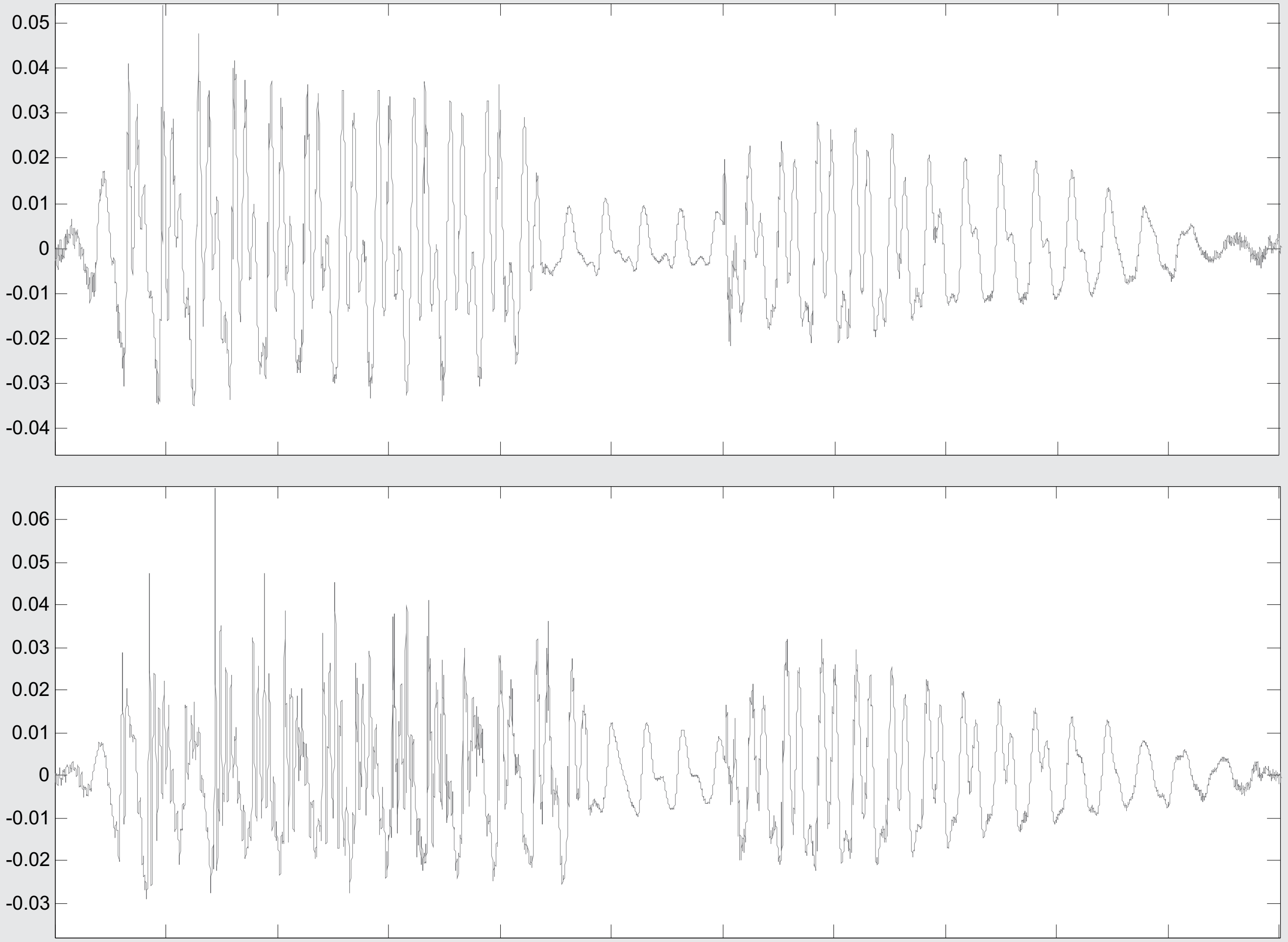
Conditional probabilities

$$p(\text{ 4 legs } | \text{ cow }) \neq p(\text{ cow } | \text{ 4 legs })$$

Bayes' Theorem
Part I:
Prior odds, likelihood ratios, & posterior odds

Given two voice recordings with acoustic properties x_1 and x_2, what is the probability that they were produced by the same speaker?

$$p(\text{ same speaker } | \text{ acoustic properties } x_1, x_2 \text{ }) = ?$$



Posterior probabilities

p(same speaker | acoustic properties x_1, x_2) = ?

p(same walker | shoe size x, shoemark size x) = ?

p(cow | x legs) = ?

Posterior odds

$$\frac{\text{p(same speaker | acoustic properties } x_1, x_2 \text{)}}{\text{p(different speaker | acoustic properties } x_1, x_2 \text{)}}$$

$$\frac{\text{p(same walker | shoe size } x\text{, shoemark size } x \text{)}}{\text{p(different walker | shoe size } x\text{, shoemark size } x \text{)}}$$

Posterior odds

$$\frac{\text{p(cow} \mid x \text{ legs)}}{\text{p(not cow} \mid x \text{ legs)}}$$

$$\frac{\text{p(} H_1 \mid E \text{)}}{\text{p(} H_2 \mid E \text{)}}$$

Likelihood ratio

$$\frac{\text{p(acoustic properties } x_1, x_2 \mid \text{same speaker)}}{\text{p(acoustic properties } x_1, x_2 \mid \text{different speaker)}}$$

$$\frac{\text{p(shoe size } x\text{, shoemark size } x \mid \text{same walker)}}{\text{p(shoe size } x\text{, shoemark size } x \mid \text{different walker)}}$$

Likelihood ratio

$$\frac{\mathrm{p}(\ x\ \mathrm{legs} \mid \mathrm{cow}\ )}{\mathrm{p}(\ x\ \mathrm{legs} \mid \mathrm{not\ cow}\ )}$$

$$\frac{\mathrm{p}(\ E \mid H_1\ )}{\mathrm{p}(\ E \mid H_2\ )}$$

Posterior odds: $$\frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

Likelihood ratio: $$\frac{p(E \mid H_1)}{p(E \mid H_2)}$$

Bayes' Theorem

initial probabilistic belief & evidence ⟶ updated probabilistic belief

higher?
or
lower?

Bayes' Theorem

initial probabilistic belief & strength of evidence → updated probabilistic belief

$$\frac{\mathrm{p}(H_1)}{\mathrm{p}(H_2)} \times \frac{\mathrm{p}(E \mid H_1)}{\mathrm{p}(E \mid H_2)} = \frac{\mathrm{p}(H_1 \mid E)}{\mathrm{p}(H_2 \mid E)}$$

prior odds ***likelihood ratio*** ***posterior odds***

Converting between probabilities and odds

$$\mathrm{p}(H_1) + \mathrm{p}(H_2) = 1$$

$$\Rightarrow \mathrm{p}(H_2) = 1 - \mathrm{p}(H_1)$$

$$\frac{\mathrm{p}(H_1)}{\mathrm{p}(H_2)} = \frac{\mathrm{p}(H_1)}{1 - \mathrm{p}(H_1)}$$

$$\mathrm{p}(H_1) = \frac{\frac{\mathrm{p}(H_1)}{\mathrm{p}(H_2)}}{1 + \frac{\mathrm{p}(H_1)}{\mathrm{p}(H_2)}}$$

- Applies to:
  - prior and posterior probabilities
  - prior and posterior odds
- Does not apply to:
  - likelihoods
  - likelihood ratios

Bayes' Theorem
Part II: Responsibilities

Bayes' Theorem

initial probabilistic belief & strength of evidence → updated probabilistic belief

$$\frac{p(H_1)}{p(H_2)} \times \frac{p(E \mid H_1)}{p(E \mid H_2)} = \frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

prior odds   ***likelihood ratio***   ***posterior odds***

Bayes' Theorem

- A forensic practitioner can**not** give the posterior odds or the posterior probability.

- For example, a forensic practitioner can**not** give the probability that two voice recordings were produced by the same speaker.

Bayes' Theorem

- Considering all the evidence presented in the case,

  determining the posterior probability of the prosecution hypothesis,

  and whether it exceeds the threshold for "beyond a reasonable doubt" or "on the balance of probabilities"

  is the task of the trier of fact (judge, panel of judges, or jury),

  not the task of the forensic practitioner.

responsibility of the trier of fact

responsibility of the trier of fact

initial probabilistic belief & strength of evidence → updated probabilistic belief

$$\frac{p(H_1)}{p(H_2)} \times \frac{p(E \mid H_1)}{p(E \mid H_2)} = \frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

prior odds — ***likelihood ratio*** — ***posterior odds***

Bayes' Theorem

- The forensic practitioner does not know what the trier of fact's prior probabilities are.
- If the forensic practitioner used their own prior probabilities, these would be either
  - arbitrary, or
  - based on knowledge of other (admissible or inadmissible) evidence in the case.

Bayes' Theorem

- The task of the forensic practitioner is to assess the strength of the particular evidence they have been asked to evaluate.

responsibility of the trier of fact

responsibility of the forensic practitioner

responsibility of the trier of fact

initial probabilistic belief & strength of evidence → updated probabilistic belief

$$\frac{p(H_1)}{p(H_2)} \times \frac{p(E \mid H_1)}{p(E \mid H_2)} = \frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

prior odds × ***likelihood ratio*** = ***posterior odds***

Bayes' Theorem
Part III: Updating beliefs

Bayes' Theorem

initial probabilistic belief	&	strength of evidence	→	updated probabilistic belief
$\frac{p(H_1)}{p(H_2)}$	×	$\frac{p(E \mid H_1)}{p(E \mid H_2)}$	=	$\frac{p(H_1 \mid E)}{p(H_2 \mid E)}$
prior odds		***likelihood ratio***		***posterior odds***

$$\frac{p(H_1)}{p(H_2)} \times \frac{p(E \mid H_1)}{p(E \mid H_2)} = \frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

responsibility of the trier of fact

responsibility of the forensic practitioner

responsibility of the trier of fact

initial probabilistic belief & strength of evidence → updated probabilistic belief

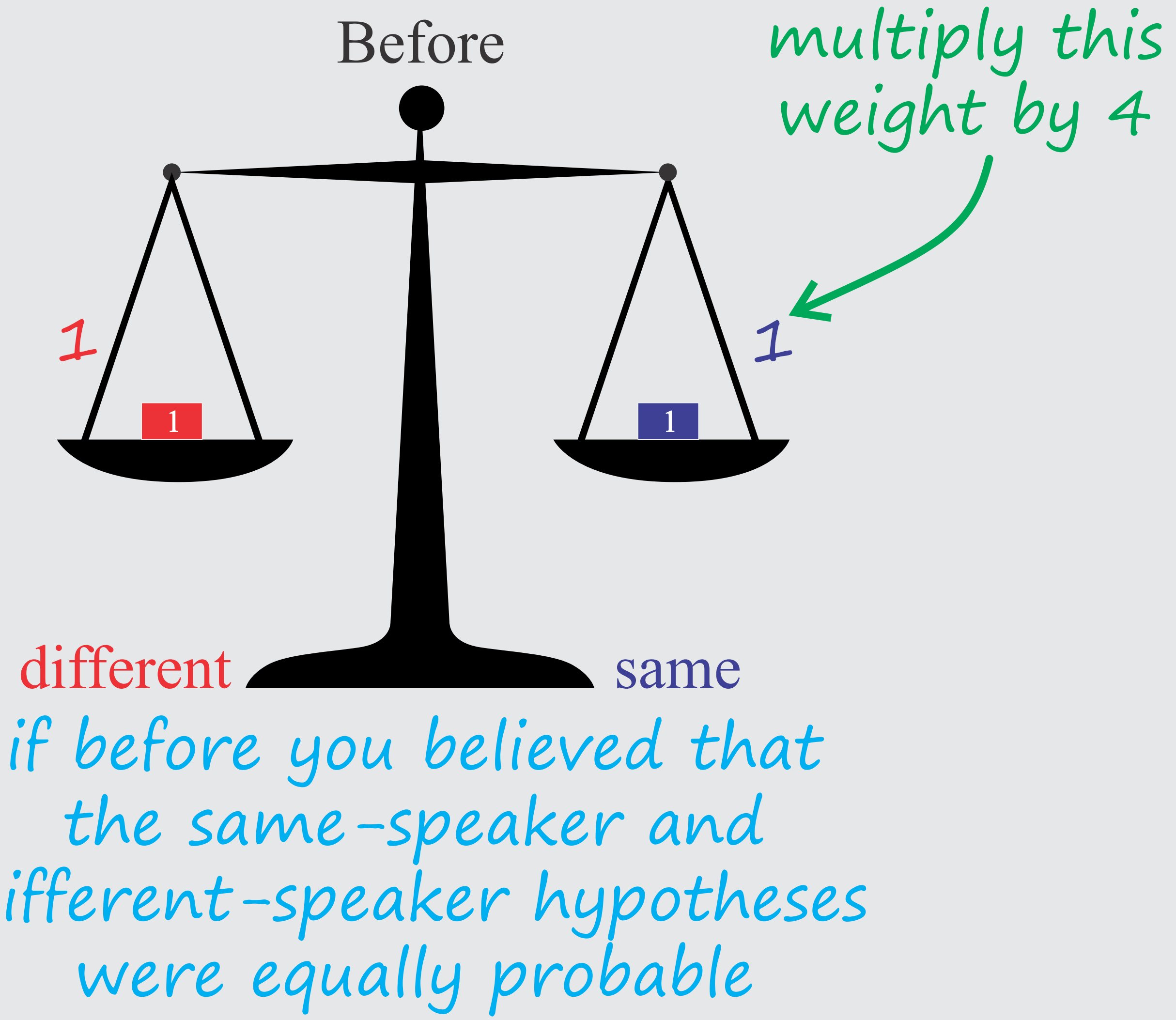
$$\frac{p(H_1)}{p(H_2)} \times \frac{p(E \mid H_1)}{p(E \mid H_2)} = \frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

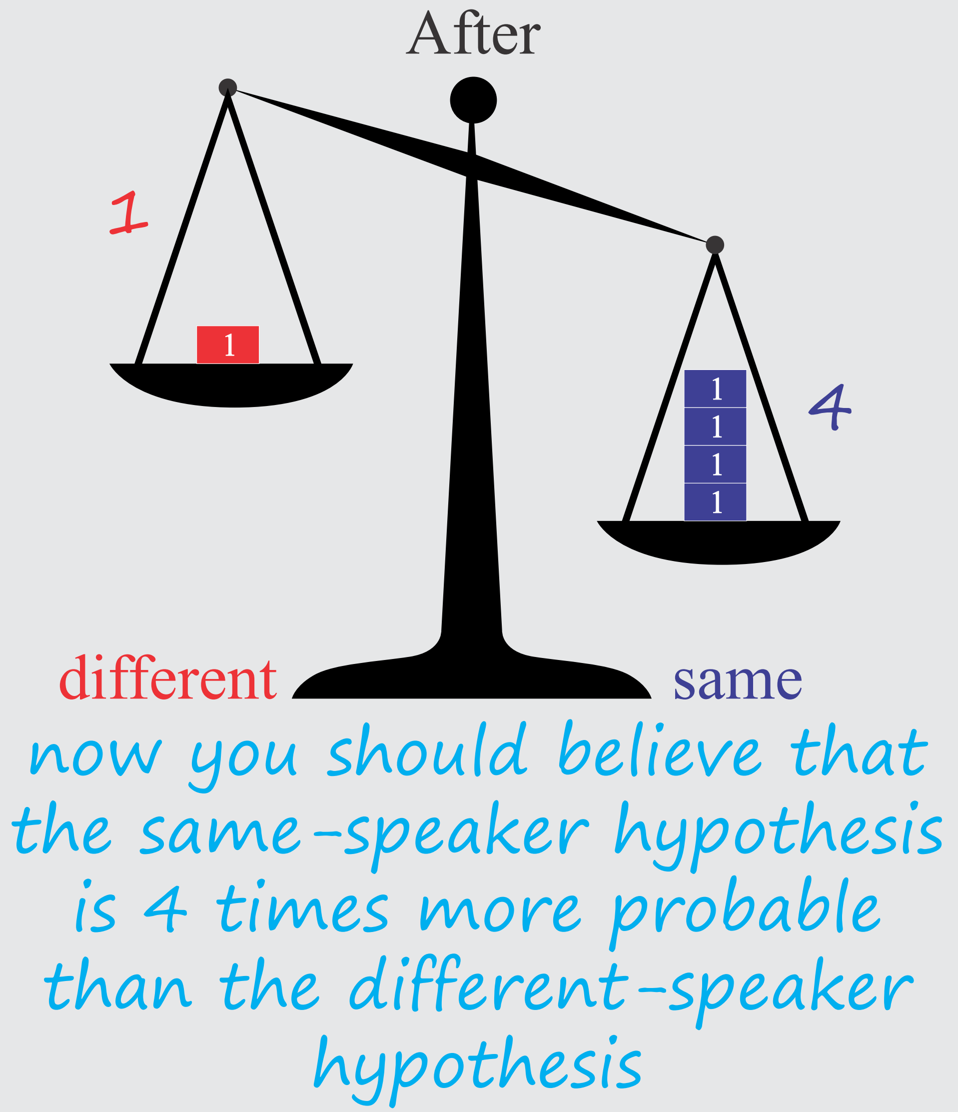
prior odds × ***likelihood ratio*** = ***posterior odds***

Likelihood ratio

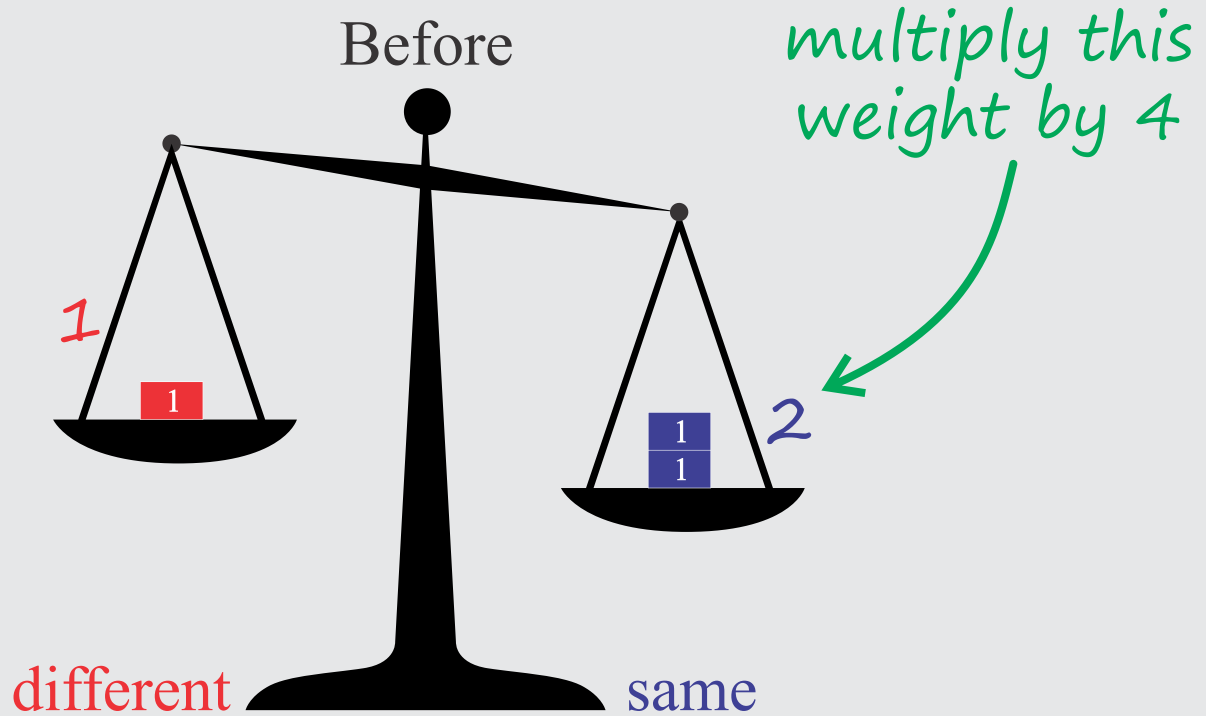
$$\frac{\text{p(acoustic properties } x_1, x_2 \mid \text{same speaker)}}{\text{p(acoustic properties } x_1, x_2 \mid \text{different speaker)}} = 4$$

The evidence is 4 times more likely given the same-speaker hypothesis than given the different-speaker hypothesis

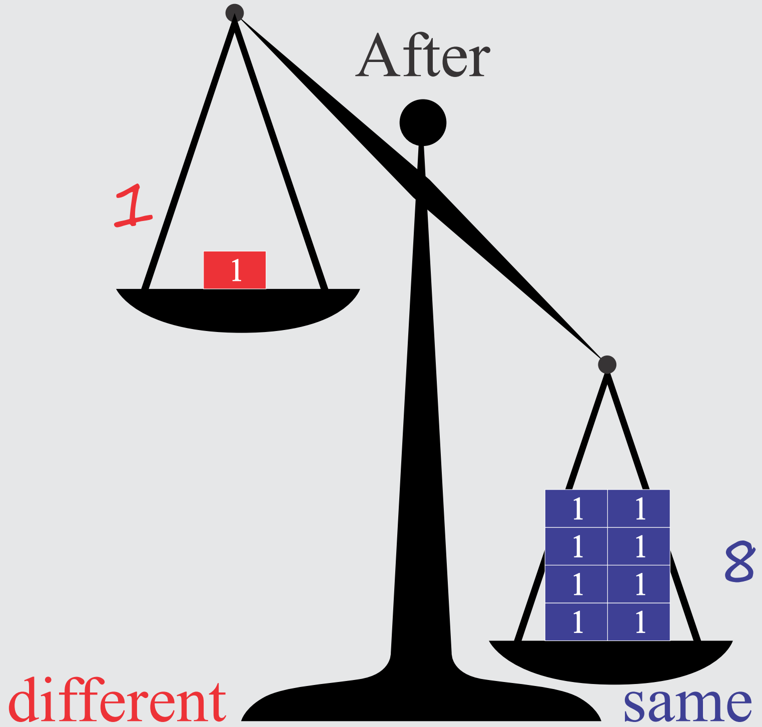




The evidence is 4 times more likely given the same-speaker hypothesis than given the different-speaker hypothesis

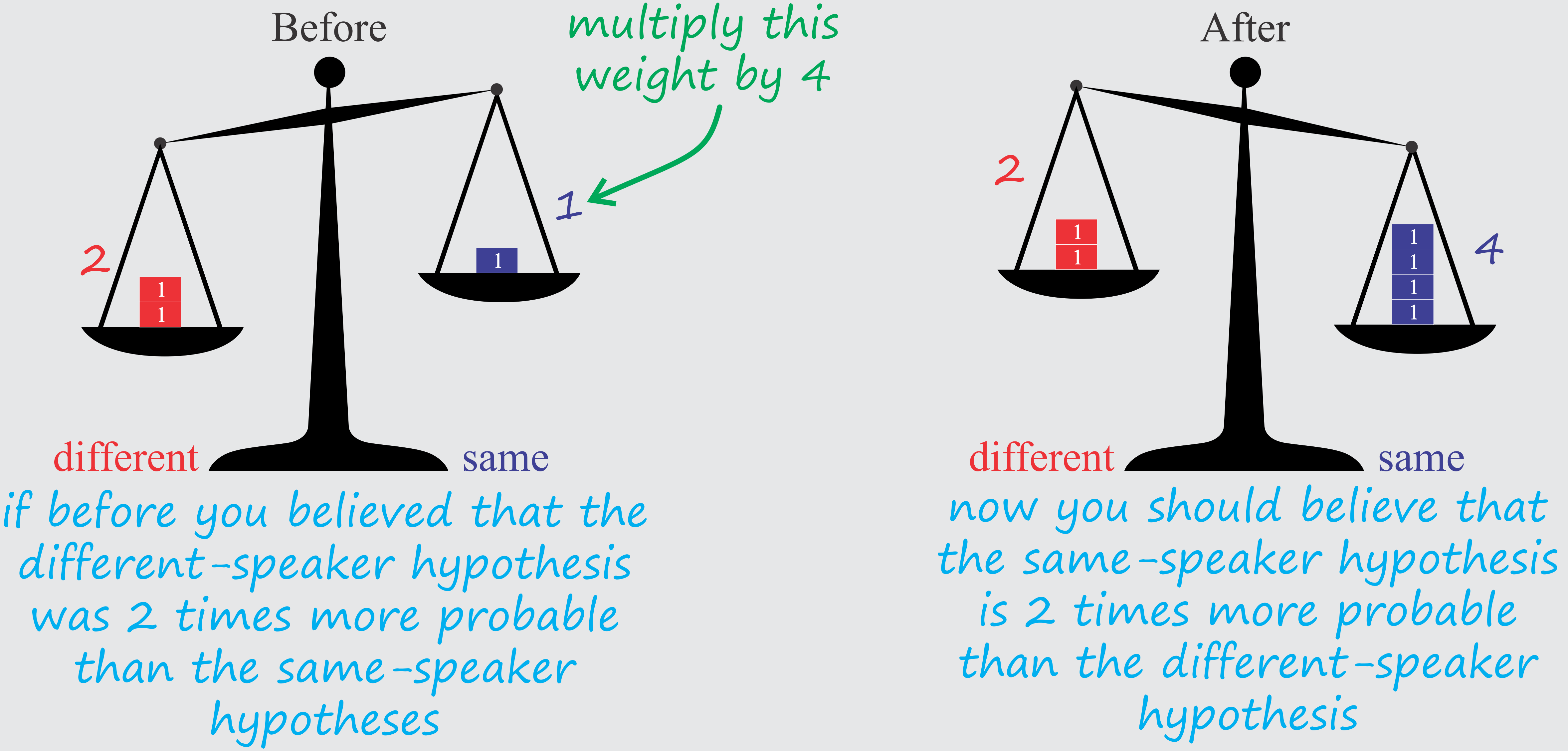


if before you believed that the same-speaker hypothesis was 2 times more probable than the different-speaker hypotheses

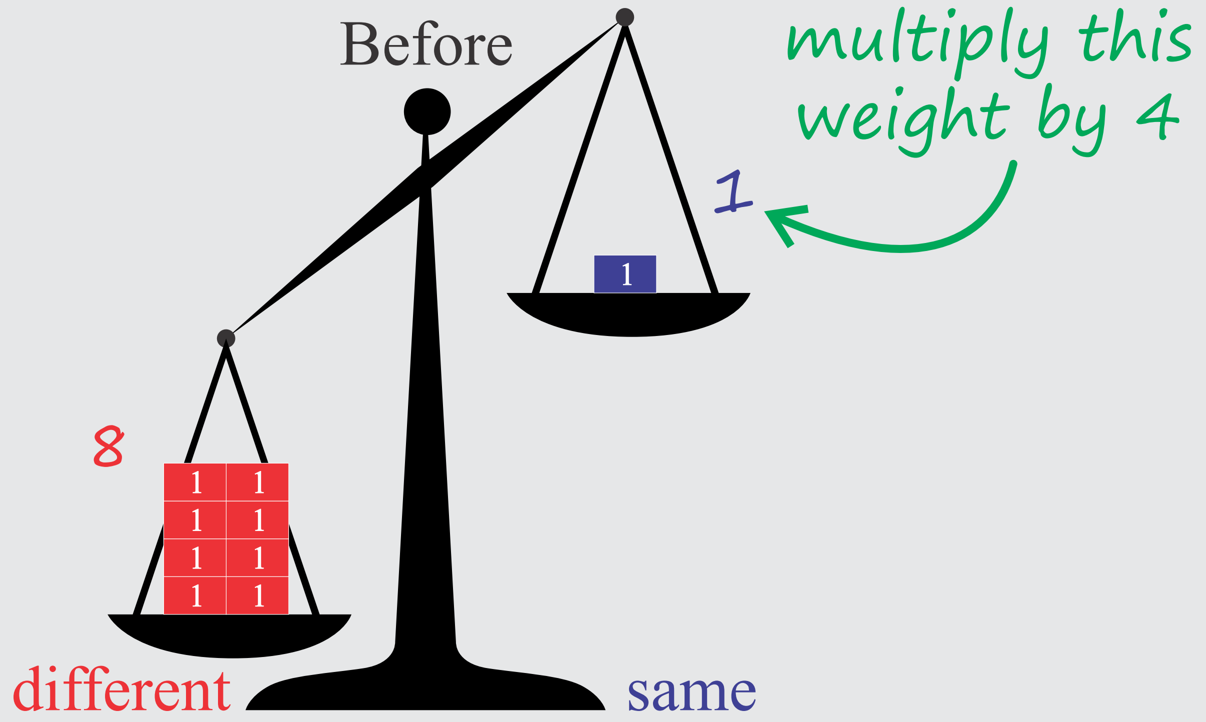


now you should believe that the same-speaker hypothesis is 8 times more probable than the different-speaker hypothesis

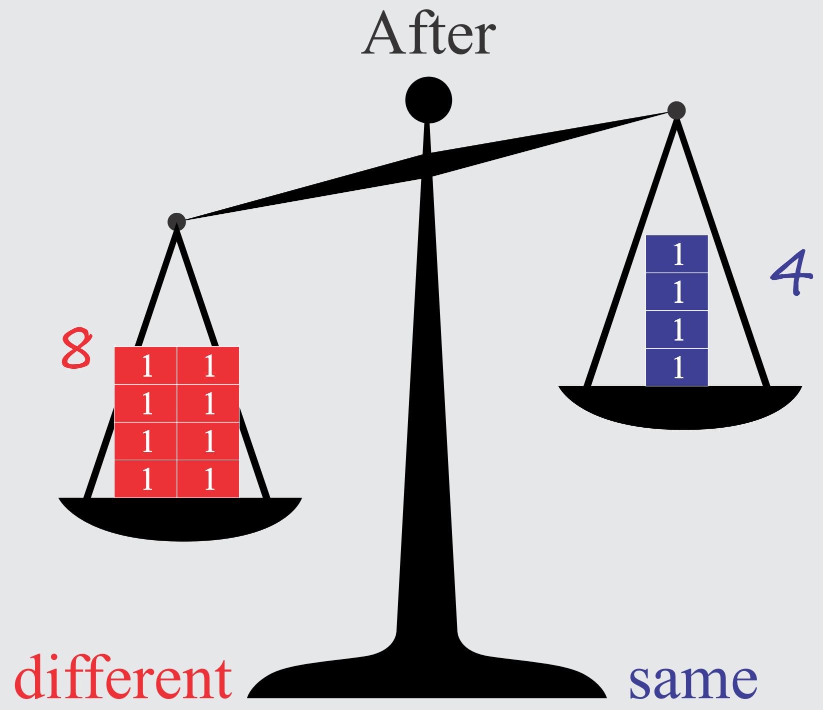
The evidence is 4 times more likely given the same-speaker hypothesis than given the different-speaker hypothesis



The evidence is 4 times more likely given the same-speaker hypothesis than given the different-speaker hypothesis



if before you believed that the different-speaker hypothesis was 8 times more probable than the same-speaker hypotheses



now you should believe that the different-speaker hypothesis is 2 times more probable than the same-speaker hypothesis

Likelihood ratio

- Work through the previous examples, but using the following likelihood-ratio values:
  - The evidence is 10 times more likely given the same-speaker hypothesis than given the different-speaker hypothesis
  - The evidence is 2 times more likely given the different-speaker hypothesis than given the same-speaker hypothesis

Illogical Reasoning

Prosecutor’s fallacy

- Forensic practitioner:

  “One would be one thousand times more likely to obtain the acoustic properties of the voice on the intercepted telephone call if it had been produced by the accused than if it had been produced by some other speaker from the relevant population.”

- Prosecutor:

  “So, to simplify for the benefit of the jury, what you are saying is that it is a thousand times more likely that the voice on the telephone intercept is the voice of the accused than the voice of any other speaker from the relevant population.”

Prosecutor's fallacy (transposed conditional)

- Forensic practitioner:

  "One would be one thousand times more likely to obtain the acoustic properties of the voice on the intercepted telephone call if it had been produced by the accused than if it had been produced by some other speaker from the relevant population."

  $$\frac{p(\,E \mid H_1\,)}{p(\,E \mid H_2\,)}$$

- Prosecutor:

  "So, to simplify for the benefit of the jury, what you are saying is that it is a thousand times more likely that the voice on the telephone intercept is the voice of the accused than the voice of any other speaker from the relevant population."

  $$\frac{p(\,H_1 \mid E\,)}{p(\,H_2 \mid E\,)}$$

What is *E*?

- From the perspective of calculating a likelihood ratio, the evidence, the ***E*** in the likelihood-ratio formula, consists of information extracted from the items of interest.

- This information will consist of quantitative measurements made of properties of items of interest or perceptual observations of properties of items of interest.

Defence attorney's fallacy

THE MATH IS CORRECT: prior odds × likelihood ratio = posterior odds

$(1/1{,}000{,}000) \times 1{,}000 = 1/1{,}000$

- Forensic practitioner:

  "One would be one thousand times more likely to obtain the properties of the fingermark had it been produced by the finger of the accused than had it been produced by a finger of some other person."

- Defence attorney:

  "Given there are approximately a million people in the region, and assuming initially that any one of them could have left the fingermark, we begin with prior odds of one over one million. Multiplying this by a likelihood ratio of one thousand, results in posterior odds of one over one thousand. Since it is one thousand times more likely that the fingermark was left by someone other than my client than that it was left by my client, this evidence fails to prove that my client left the finger mark. Therefore, it should not be taken into consideration by the jury."

Defence attorney’s fallacy

THE MATH IS CORRECT: prior odds × likelihood ratio = posterior odds

$(1/1{,}000{,}000) \times 1{,}000 = 1/1{,}000$

- Forensic practitioner:

  “One would be one thousand times more likely to obtain the properties of the fingermark had it been produced by the finger of the accused than had it been produced by a finger of some other person.”

- Defence attorney:

  “Given there are approximately a million people in the region, and assuming initially that any one of them could have left the fingermark, we begin with prior odds of one over one million. Multiplying this by a likelihood ratio of one thousand, results in posterior odds of one over one thousand. Since it is one thousand times more likely that the fingermark was left by someone other than my client than that it was left by my client, this evidence fails to prove that my client left the finger mark. Therefore, it should not be taken into consideration by the jury.”

Defence attorney's fallacy (small-number fallacy)

$$\frac{p(H_1)}{p(H_2)} \times \frac{p(E \mid H_1)}{p(E \mid H_2)} = \frac{p(H_1 \mid E)}{p(H_2 \mid E)}$$

evidence type	prior odds	likelihood ratio	posterior odds
fingerprints	1/1,000,000	1,000	1/1,000
footwear	1/1,000	1,000	1
voice recordings	1	1,000	1,000

Trier of fact's fallacy

- Forensic practitioner:

  "One would be one billion times more likely to obtain the properties of the DNA found at the crime scene had it come from the accused than had it come from some other person in the country."

- Trier of fact:

  "One billion is a very large number. The DNA must have come from the accused. I can ignore other evidence which suggests that it did not come from the accused."

Trier of fact's fallacy (large-number fallacy)

- Forensic practitioner:

  "One would be one billion times more likely to obtain the properties of the DNA found at the crime scene had it come from the accused than had it come from some other person in the country."

- Trier of fact:

  "One billion is a very large number. The DNA must have come from the accused. I can ignore other evidence which suggests that it did not come from the accused."

Investigator’s fallacy

- Forensic practitioner:
  “One would be one hundred times more likely to obtain the properties of the glass fragments found on the suspect’s clothing had they come from the broken window than had they come from some other window in the region.”

- Investigator:
  “My belief that the suspect broke the window versus that someone else broke the window is 100 time greater than it was before I got the forensic practitioner’s report.”

Investigator's fallacy (change of level on hierarchy of propositions)

- Forensic practitioner:

  "One would be one hundred times more likely to obtain the properties of the glass fragments found on the suspect's clothing had they come from the broken window than had they come from some other window in the region."

  source-level propositions

- Investigator:

  "My belief that the suspect broke the window versus that someone else broke the window is 100 time greater than it was before I got the forensic practitioner's report."

  activity-level propositions

Hierarchy of propositions

- **Offense level:**
  - the suspect committed the robbery
- **Activity level:**
  - the suspect broke the window
- **Source level:**
  - the glass fragments came from the broken window

Thank Moo