Introduction to logical reasoning for the evaluation of forensic evidence

Geoffrey Stewart Morrison

\[
\frac{p(E|H_p)}{p(E|H_d)}
\]
Concerns

- Logically correct framework for evaluation of forensic evidence
  - ENFSI Guideline for Evaluative Reporting 2015

- But what is the warrant for the opinion expressed? Where do the numbers come from?
  - R v T 2010; Risinger at ICFIS 2011

- Demonstrate validity and reliability

- Transparency
  - R v T 2010

- Reduce potential for cognitive bias
  - NIST/NIJ Fingerprint analysis 2012; NCFS task-relevant information 2015

- Communicate strength of forensic evidence to triers of fact
Paradigm

- **Use of the likelihood-ratio framework for the evaluation of forensic evidence**
  - logically correct

- **Use of relevant data (data representative of the relevant population), quantitative measurements, and statistical models**
  - transparent and replicable
  - resistant to cognitive bias

- **Empirical testing of validity and reliability under conditions reflecting those of the case under investigation, using test data drawn from the relevant population**
  - only way to know how well it works
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?
• This 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
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 are a forensic shoe comparison expert...

suspect’s shoe

crime-scene footprint
Imagine you are a forensic shoe comparison expert...

• The footprint at the crime scene is size 10

• The suspect’s shoe is size 10
  – What is the probability of the footprint at the crime scene would be size 10 if it had been made by the suspect’s shoe? (similarity)

• Half the shoes at the recycling depot are size 10
  – What is the probability of the footprint at the crime scene would be size 10 if it had been made by the someone else’s shoe? (typicality)
Imagine you are a forensic shoe comparison expert...

- The footprint at the crime scene is size 14

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

- 1% of the shoes at the recycling depot are size 14
  - What is the probability of the footprint at the crime scene would be size 14 if it had been made by the someone else’s shoe? (typicality)
Imagine you are a forensic shoe comparison expert...

- The footprint at the crime scene and the suspect’s shoe are both size 10

\[
\frac{\text{similarity}}{\text{typicality}} = \frac{1}{0.5} = 2
\]

you are twice as likely to get a size 10 footprint 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 are a forensic shoe comparison expert...

- The footprint at the crime scene and the suspect’s shoe are both size 14

\[
\text{similarity} / \text{typicality} = 1 / 0.01 = 100
\]

you are 100 times more likely to get a size 14 footprint at the crime scene \textit{if it were produced by the suspect’s shoe} than \textit{if it were produced by someone else’s shoe}

- someone else selected at random from the relevant population
Imagine you are a forensic shoe comparison expert...

- size 10
  \[
  \text{similarity} / \text{typicality} = \frac{1}{0.5} = 2
  \]

- size 14
  \[
  \text{similarity} / \text{typicality} = \frac{1}{0.01} = 100
  \]

- If you didn’t have a database, could you have made subjective estimates at relative proportions of different shoe sizes in the population and applied the same logic to arrive at a conceptually similar answer?
¿Area?
similarity / typicality = likelihood ratio
Given that it is a cow, what is the probability that it has four legs?

\[ p(4 \text{ legs} \mid \text{cow}) = ? \]
Given that it has four legs, what is the probability that it is a cow?

\[ p( \text{cow} \mid 4 \text{ legs} ) = ? \]
Given two voice samples with acoustic properties $x_1$ and $x_2$, what is the probability that they were produced by the same speaker?

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

\[ p( \text{same wearer} \mid \text{shoe size } x, \text{footprint size } x ) = ? \]

\[ p( \text{cow} \mid x \text{ legs} ) = ? \]
Bayes’ Theorem

\[
\text{posterior odds} = \frac{p(\text{same speaker} | \text{acoustic properties } x_1, x_2)}{p(\text{different speaker} | \text{acoustic properties } x_1, x_2)}
\]

\[
= \frac{p(\text{acoustic properties } x_1, x_2 | \text{same speaker})}{p(\text{acoustic properties } x_1, x_2 | \text{different speaker})} \times \frac{p(\text{same speaker})}{p(\text{different speaker})}
\]

\text{likelihood ratio}

\text{prior odds}
Bayes’ Theorem

initial probabilistic belief + evidence → updated probabilistic belief

higher? or lower?
However !!!

The forensic scientist acting as an expert witness can **NOT** give the posterior probability. They can **NOT** give the probability that two speech samples were produced by the same speaker.
Why not?

- The forensic scientist does not know the prior probabilities.

- Considering all the evidence presented so as to determine the posterior probability of the prosecution hypothesis and whether it is true 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 scientist.

- The task of the forensic scientist is to present the strength of evidence with respect to the particular samples provided to them for analysis. They should not consider other evidence or information extraneous to their task.
Bayes’ Theorem

\[
\text{posterior odds} \quad \frac{p(\text{same speaker} \mid \text{acoustic properties } x_1, x_2)}{p(\text{different speaker} \mid \text{acoustic properties } x_1, x_2)} = \frac{p(\text{acoustic properties } x_1, x_2 \mid \text{same speaker})}{p(\text{acoustic properties } x_1, x_2 \mid \text{different speaker})} \times \frac{p(\text{same speaker})}{p(\text{different speaker})}
\]

\text{likelihood ratio}
Bayes’ Theorem

\[
p(\text{same speaker} \mid \text{acoustic properties } x_1, x_2) \quad \frac{p(\text{different speaker} \mid \text{acoustic properties } x_1, x_2)}{p(\text{same speaker} \mid \text{acoustic properties } x_1, x_2)} = \frac{p(\text{acoustic properties } x_1, x_2 \mid \text{same speaker})}{p(\text{acoustic properties } x_1, x_2 \mid \text{different speaker})} \times \frac{p(\text{same speaker})}{p(\text{different speaker})}
\]
Bayes’ Theorem

\[
\text{posterior odds} = \frac{p(\text{same speaker} \mid \text{acoustic properties } x_1, x_2)}{p(\text{different speaker} \mid \text{acoustic properties } x_1, x_2)}
\]

\[
= \frac{p(\text{acoustic properties } x_1, x_2 \mid \text{same speaker})}{p(\text{acoustic properties } x_1, x_2 \mid \text{different speaker})} \times \frac{p(\text{same speaker})}{p(\text{different speaker})}
\]

\text{likelihood ratio}

\text{prior odds}
Likelihood Ratio

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

\[
\frac{p(\text{shoe size } x, \text{footprint size } x \mid \text{same wearer})}{p(\text{shoe size } x, \text{footprint size } x \mid \text{different wearer})}
\]

\[
\frac{p(\text{x legs} \mid \text{cow})}{p(\text{x legs} \mid \text{not a cow})}
\]

\[
\frac{p(\text{E} \mid H_{\text{prosecution}})}{p(\text{E} \mid H_{\text{defence}})}
\]
Example

- Forensic scientist: You would be **4 times more likely** to get the acoustic properties of the voice on the offender recording if it were produced by the suspect than if it were produced by some other speaker selected at random from the relevant population.

- Whatever the trier of fact’s belief as to the relative probabilities of the same-speaker versus the different-speaker hypotheses before being presented with the likelihood ratio, after being presented with the likelihood ratio their relative belief in the probability that the voices on the two recordings belong to the same speaker versus the probability that they belong to different speakers should be **4 times greater than it was before.**
The evidence is 4 time more likely given the same-speaker hypothesis than given the different-speaker hypothesis.

Before

If before you believed that the same-speaker and different-speaker hypotheses were equally probable, multiply this weight by 4.

After

Now you should believe that the same-speaker hypothesis is 4 times more probable than the different-speaker hypothesis.
The evidence is 4 time more likely given the same-speaker hypothesis than given the different-speaker hypothesis.

Before

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

After

now you should believe that the same-speaker hypothesis is 8 times more probable than the different-speaker hypothesis.
The evidence is 4 time more likely given the same-speaker hypothesis than given the different-speaker hypothesis.

Before

- different: 2
- same: 1

After

- different: 2
- same: 4

If before you believed that the different-speaker hypothesis was 2 times more probable than the same-speaker hypotheses, now you should believe that the same-speaker hypothesis is 2 times more probable than the different-speaker hypothesis.
The evidence is 4 time more likely given the same-speaker hypothesis than given the different-speaker hypothesis.

Before

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

multiply this weight by 4

After

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

discrete data
Ready to calculate a likelihood ratio?

\[
\frac{p(E \mid H_{\text{prosecution}})}{p(E \mid H_{\text{defence}})}
\]

\[
\frac{p(x \text{ legs} \mid \text{cow})}{p(x \text{ legs} \mid \text{not a cow})}
\]
Discrete data: bar graph

- Cows
- Not cows

Proportion of animals with different numbers of legs.
\[ \frac{p(4 \text{ legs} \mid \text{cow})}{p(4 \text{ legs} \mid \text{not a cow})} \]

- **cows**
- **not cows**

Proportion

- 0.98 →
- 0.49 ←
\[
\frac{p(\text{4 legs | cow})}{p(\text{4 legs | not a cow})} = 2
\]
Relevant Population
Imagine you are a forensic hair comparison expert...

- The hair at the crime scene is blond
- The suspect has blond hair
- What do you do?
Imagine you are a forensic hair comparison expert...

- The hair at the crime scene is blond
- The suspect has blond hair

\[
p( \text{blond hair at crime scene} \mid \text{suspect is source}) \]
\[
p( \text{blond hair at crime scene} \mid \text{someone else is source})
\]
Imagine you are a forensic hair comparison expert...

- The hair at the crime scene is blond
- The suspect has blond hair

\[
p(\text{blond hair at crime scene} \mid \text{suspect is source})
\]

\[
p(\text{blond hair at crime scene} \mid \text{someone else is source})
\]

- Someone else selected at random from the relevant population
Imagine you are a forensic hair comparison expert...

- The hair at the crime scene is blond
- The suspect has blond hair

\[
p(\text{blond hair at crime scene} \mid \text{suspect is source})
\]

\[
p(\text{blond hair at crime scene} \mid \text{someone else is source})
\]

- Someone else selected at random from the relevant population
- What is the relevant population?
Imagine you are a forensic hair comparison expert...

- The hair at the crime scene is blond
- The suspect has blond hair

\[
p(\text{blond hair at crime scene} \mid \text{suspect is source})
\]

\[
p(\text{blond hair at crime scene} \mid \text{someone else is source})
\]

- Someone else selected at random from the relevant population
- What is the relevant population?
  - Stockholm
  - Beijing
- You need to use a sample representative of the relevant population
A likelihood ratio is the answer to a **specific question** defined by the prosecution and the defence hypotheses.

The **defence** hypothesis specifies the **relevant population**.

The forensic scientist must make explicit the specific question they answered so that the trier of fact can:

– understand the question
– consider whether the question is an appropriate question
– understand the answer
Fallacies of Interpretation
Prosecutor’s Fallacy

- Forensic Scientist:
  - One would be one thousand times more likely to obtain the acoustic properties of the voice on the intercepted telephone call had it 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 if I may, 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 (transposition of the conditionals)

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

\[
\frac{P(E | H_{\text{prosecution}})}{P(E | H_{\text{defence}})}
\]

- Prosecutor:
  - So, to simplify for the benefit of the jury if I may, 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.

\[
\frac{P(H_{\text{prosecution}} | E)}{P(H_{\text{defence}} | E)}
\]
Defence Attorney’s Fallacy (small number fallacy)

- Forensic Scientist:
  - One would be one thousand times more likely to obtain the measured properties of the partial latent finger mark had it been produced by the finger of the accused than if it had been produced by a finger of some other person.

- Defence Attorney:
  - So, given that there are approximately a million people in the region and assuming initially that any one of them could have left the finger mark, we begin with prior odds of one over one million, and the evidence which has been presented has resulted in posterior odds of one over one thousand. One over one thousand is a small number. Since it is one thousand times more likely that the finger mark was left by someone other than my client than that it was left by my client, I submit that this evidence fails to prove that my client left the finger mark and as such it should not be taken into consideration by the jury.

THE MATHEMATICS IS CORRECT: \[
\text{prior odds} \times \text{likelihood ratio} = \text{posterior odds} \]

\[
\frac{1}{1,000,000} \times 1,000 = \frac{1}{1,000}
\]
Trier of Fact’s Fallacy  (large number fallacy)

• Forensic Scientist:
  – One would be one billion times more likely to obtain the measured properties of the DNA sample from 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 sample must have come from the accused. I can ignore other evidence which suggests that it did not come from him.
Likelihood Ratio Calculation II

continuous data
Discrete data: bar graph

\[
\frac{p(4 \text{ legs} \mid \text{cow})}{p(4 \text{ legs} \mid \text{not a cow})} = 2
\]
Continuous data: histograms → probability density functions (PDFs)
Continuous data: histograms $\rightarrow$ probability density functions (PDFs)
The graph shows two probability density models: the "suspect model" and the "population model". The suspect model is a higher and sharper peak compared to the population model. The x-axis represents the variable x, and the y-axis represents the probability density.
\[ LR = \frac{0.021}{0.005} = 4.02 \]

The diagram illustrates the probability density functions for the suspect model and the population model. The suspect model has a higher probability density at the offender value, indicating a higher likelihood for the suspect.
Past, Present, Future
Likelihood ratios

- 1906 retrial of Alfred Dreyfus
- Jean-Gaston Darboux, Paul Émile Appell, Jules Henri Poincaré
Likelihood ratios

- Adopted as standard for evaluation of DNA evidence in mid 1990’s
Likelihood ratios

- Association of Forensic Science Providers (2009)
  - Standards for the formulation of evaluative forensic science expert opinion

- 31 signatories [from Aitken to Zadora] (2011)
  - Expressing evaluative opinions: A position statement

  - Guideline for evaluative reporting in forensic science

- President’s Council of Advisors on Science & Technology (2016)
  - Ensuring scientific validity of feature-comparison methods
Thank You

http://geoff-morrison.net/

http://forensic-evaluation.net/