Forensic interpretation and evaluation of evidence

...according to the model used at The Swedish National Forensic centre

...and with some considerations for medico-legal casework

Special lecture within the course Decision Theory, fall semester 2016
Evidence evaluation…

… is an *inductive* inferential process.

”Draw conclusions about what has happened from observed consequences of what happened – what we observe afterwards”

Packages of amphetamine was delivered in a car. ➔ Traces of amphetamine recovered from a seat in the car.

?
The two modes of the forensic process

- Investigative mode
  - Formulation of hypotheses regarding the activities (that have taken place at the crime scene)
  - Definition of criteria for subsequent recovery of traces

- Evaluative mode
  - Specific questions (formulated hypotheses) about recovered traces and their possible links to suspects or seized goods are treated by use of probabilistic reasoning

...but the two modes come interchangeably in course of the investigation:
More about the investigative mode

• In the investigative mode *several hypotheses are formulated* about what might have happened at a crime scene, a scene of fire, a finding-place,…

• Assessment and interpretation of detailed observations lead up to a ranking of the hypotheses formulated

• The assessment is made by deeming how expected detailed observations are under each of the hypotheses formulated. – and falsify such hypotheses under which the observations are considered improbable

  This is where the evaluative mode may enter…

• The finally retained hypotheses would form a context that serve as the explanation delivered about what happened at the scene – a kind of giant hypothesis, supported by the investigation, but to be challenged in court
The evaluative mode

• Some questions (hypotheses) cannot be assessed – completely or directly – at the scene (of crime, of fire)
• E.g. hypotheses linking recovered traces to subsequently identified suspects or seized material
  – Was it this shoe that made the recovered footwear mark?
  – Does the blood on the floor originate from the dead person found in the villa?
  – Were the glass fragments recovered from the suspect’s jacket transferred from the smashed window at the crime scene?
  – Did somebody burn hazardous waste here?
  – Were the two seized bags of amphetamine parts of the same manufacturing batch?
  – …

Common for these types of questions is that they have the ultimate answers Yes and No.
• Conclusions from the evaluative mode may

- Sort things out at a crime scene/finding-place that helps in deciding along which path the subsequent investigation should continue…

  … should the conclusion be interpreted as a Yes or a No by the CSI

- Be used as support for an hypothesis about a certain course at the crime scene/finding-place

- constitute a self-standing piece of forensic evidence that links a recovered material to an individual or another control material, or a specific class/category

Is it then possible to always conclude with a Yes or No?
Source level attributions

A recovered footwear mark and a pair of shoes seized with a suspect.

Blood recovered from a garment and DNA from swabbing a suspect.

Two seizures of amphetamine – same origin?
Forensic investigation – and evaluation…

Two seizures of amphetamine – same origin?

Findings:
- The two seizures show similarities in their impurity profiles (presence of small amounts of other substances than amphetamine – bi-products in the manufacturing)
- The dry concentrations differ between the seizures

What do these findings signify?
Who requested the forensic investigation?

The leader of the preliminary investigation (the prosecutor)

What kind of answer is (ideally) expected?

”The two seizures have a common origin” (Yes)

or

”The two seizures have different origins” (No)

Is it possible to answer Yes?
The question (most probably) put by the commissioner…

Do the two seizures have a common origin?

…can be re-formulated into a main hypothesis

\[ H_m : \text{The two seizures have a common origin} \]

The main hypothesis is a statement that constitutes one explanation – but not necessarily a good one – to the findings obtained.

- Similarities between impurity profiles
- Different dry concentrations

The weight as evidence of the statement consists of the belief in the statement – and its relevance for the current alleged activity.

\textit{N.B!} \( H_m \) can only be true or false. It is the uncertainty about its truth that is the subject of discussion.
Focusing on the belief (or what would be a proper expression)...

When the veracity is neither 100% nor 0% with certainty we must use probability calculus.

Is it actually possible to directly estimate the probability that

\[ H_m : \text{The two seizures have a common origin} \]

is true?

Answer: No.
This probability is deemed on by combining the forensic evaluation with other (non-forensic) information from the investigation (supporting or non-supporting \( H_m \)).
$H_m$: The two seizures have a common origin

Final probability of $H_m$ being true
Alternative hypothesis

\[ H_a : \text{The two seizures have different origins} \]

Should be chosen to cover all relevant alternatives to the main hypothesis.
Forensic evaluation

How expected/probable are…

- similarities between impurity profiles
- different dry concentrations

…if the main hypothesis is true?  \( \Rightarrow P(\text{Findings} \mid H_m) \)

…if the alternative hypothesis is true?  \( \Rightarrow P(\text{Findings} \mid H_a) \)

Forensic value of evidence = \( V = \frac{P(\text{Findings} \mid H_m)}{P(\text{Findings} \mid H_a)} \)

\( V > 1 \Rightarrow \) The findings are \( V \) times more probable if \( H_m \) is true compared to if \( H_a \) is true

\( V < 1 \Rightarrow \) The findings are \( 1/V \) times more probable if \( H_a \) is true compared to if \( H_m \) is true
How probable – prior to the forensic investigation – is...

\( H_m : \) The two seizures have a common origin

\[ \Rightarrow P(H_m) \]

...and how probable – prior to the forensic investigation – is ...

\( H_a : \) The two seizures have different origins

\[ \Rightarrow P(H_a) \]

Prior odds = \( O = \frac{P(H_h)}{P(H_a)} \)
Other information

\[ O = \frac{P(H_h)}{P(H_a)} \]

\[ V = \frac{P(\text{Findings} \mid H_h)}{P(\text{Findings} \mid H_a)} \]

Bayes' theorem:

\[ \frac{P(H_h \mid \text{Findings})}{P(H_a \mid \text{Findings})} = V \times O \]

\[ P(H_h \mid \text{Findings}) \]

Final probability of \( H_m \) being true-

Weight as evidence (veracity)
Weight as evidence for whom?...

A source level attribution is in general a forensic investigation with the prosecutor as ”destination”.

It is the prosecutor (via the police leader of the preliminary investigation) who (at least in theory) …

- is involved with formulating the alternative hypothesis
- has to deem on the magnitude of the prior odds
- must consider if…

...the weight as evidence is sufficient to bring this hypothesis as evidence to the indictment
Bayes’ theorem both in terms of a mathematical formula and as a graphical description

\[
\frac{P(H_h)}{P(H_a)} \times \frac{P(E|H_h)}{P(E|H_a)} = \frac{P(H_h|E)}{P(H_a|E)}
\]

**Prior odds**
Measures how certain/uncertain the commissioner (prosecutor, police, judge) is about the truth of \( H_m \) before considering the outcome of the forensic investigation.

**Forensic value of evidence, \( V \)**
States how much more (or less) probable the forensic findings are if \( H_m \) is true compared to if \( H_a \) is true.

Commonly used term: **Likelihood ratio**

**Posterior odds**
Measures how certain/uncertain the commissioner is about the truth of \( H_m \) upon considering the outcome of the forensic investigation

\( \Rightarrow \) **Weight as evidence**
Prior odds

\[
\frac{P(H_h)}{P(H_a)} \times \frac{P(E|H_h)}{P(E|H_a)} = \frac{P(H_h|E)}{P(H_a|E)}
\]

Posterior odds

Forensic value of evidence, \( V \)
• With the same pair of main and alternative hypothesis the forensic findings always have the same forensic value of evidence (likelihood ratio) i.e. the arrow has the same angle.

• The weights as evidence (the posterior odds) can on the other hand differ depending on the magnitude of the prior odds.
Estimation/calculation – in practice – of the magnitude of the forensic value of evidence ($V$)

In most forensic subject fields today there are no validated mathematical models to support the calculation of the forensic values of evidence.

Lack of background/reference data is the main explanation.

A forensic laboratory should however have uniform standards for reporting their values of evidence.

When models are lacking the components of the value of evidence (i.e. the probabilities in the numerator and denominator of $V$) must be assigned based on (subjective and/or collective) experience and subject knowledge.

$\Rightarrow$ Fairly rough estimates of the magnitudes

$\Rightarrow$ All reporting of evidence from NFC – with or without using data bases and/or mathematical models – are made using a common ordinal scale of conclusions!
The scale of conclusion used at NFC:

<table>
<thead>
<tr>
<th>level</th>
<th>Magnitude of $V$</th>
<th>&quot;Explanation&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>+4</td>
<td>at least one million</td>
<td>The findings are…at least one million times more expected…</td>
</tr>
<tr>
<td>+3</td>
<td>between 6000 and one million</td>
<td>…at least 6000 times more expected…</td>
</tr>
<tr>
<td>+2</td>
<td>between 100 and 6000</td>
<td>…at least 100 times more expected…</td>
</tr>
<tr>
<td>+1</td>
<td>between 6 and 100</td>
<td>…at least 6 times more expected…</td>
</tr>
<tr>
<td>0</td>
<td>between 1/6 and 6</td>
<td>… equally expected …</td>
</tr>
<tr>
<td>−1</td>
<td>between 1/100 and 1/6</td>
<td>…if the main hypothesis is true compared to if the alternative hypothesis is true</td>
</tr>
<tr>
<td>−2</td>
<td>between 1/6000 and 1/100</td>
<td>…at least 6 times more expected…</td>
</tr>
<tr>
<td>−3</td>
<td>between 1/(one million) and 1/6000</td>
<td>…at least 6000 times more expected…</td>
</tr>
<tr>
<td>−4</td>
<td>at most 1/(one million)</td>
<td>…at least one million times more expected…</td>
</tr>
</tbody>
</table>

…if the alternative hypothesis is true compared to if the main hypothesis is true
Is this something specific for NFC Sweden?

**Paternity index** (Gürtler (1956)) – Early introduction of the *Likelihood Ratio*

Lindley (1977) ”A problem in forensic science”

Forensic Science Service, England and Wales: ”Case Assessment and Interpretation”, 1990s

Ecole des Sciences Criminelles, Université de Lausanne, CH

LGC Forensics Ltd, England

ESR, New Zealand

NFI, The Netherlands

Guardia Civil, Spain

NFC (SKL), Sweden

INCC, Belgium

IES, Krakow, Poland

EFE, Ireland
Hypotheses at activity level

Were the shoes worn by someone who was in the flat when the crime was committed?

Was the suspect in physical contact with the garment?

Was the violence done from the front against the head?
More interesting to consider your daily casework...

Assume the injured party states: “My husband hit me with a stick several times in the face and gave me a number of kicks in my back with his feet”

Could this constitute the main hypothesis of the case?

What would be a relevant alternative hypothesis?

- All other explanations to the characteristics of the wounds observed in the medical examination?
- The story of the husband (if any?), for instance “She stumbled and fell in the stairs”
Assume we use the framework with hypotheses

\[ H_m \] The injured party was exposed to blunt force trauma in her face and…”

When writing a statement using a probability scale do you actually address the probability of \( H_m \)?

…or do you address the probability of your findings upon examination under the preliminary assumption that any alternative to \( H_m \) holds?
Is it…

\( H_m \) The injured party was exposed to blunt force trauma in her face and…”

that with a certain probability implies

\( E \): The observed characteristics of the wounds

\[ H_m \implies E \]

…or is it…

\( E \): The observed characteristics of the wounds

that with a certain probability implies

\( H_m \) The injured party was exposed to blunt force trauma in her face and…”

\[ E \implies H_m \]
It is good practice to – in a logical and coherent way

- separate between the hypotheses of a case and the findings from the examination
- take into account the (im)balance between the two hypotheses – prior and posterior to the examination
How were the levels of the NFC scale derived?

• The forensic value of evidence is a ratio of two probabilities

• In theory this value can vary from
  – zero (exclusion of the *main hypothesis*)
  to
  – infinity (exclusion of the *alternative hypothesis*)

• …but for a scale to be useful in practice the number of levels must be limited – decomposition of an infinitely long interval into a finite number of intervals

• We chose a set of levels *symmetrically* spread around the value 1, with four supporting levels (positive) and four non-supporting levels (negative), each level corresponding to an interval of values

• The lower limits of the intervals on the supporting side were chosen so that the final weight as evidence (posterior probability) would be acceptably high with respect to each level when the prior odds are equal to one.
Grad + 2: Sannolikheten 0,99 (= 99 %) är i rättssammanhang allmänt vedertagen som tillräckligt hög för att ”styrka” en hypotes (att ett enskilt resultat tas med i bevisningen)

Grad + 4: Sannolikheten 0,999999 är ett historiskt ”arv” från hur +4 har valts för resultatvärden för dna-överensstämmelser

Grad +1 och + 3: Intervallen av resultatvärden har valts så att de successivt ökar i längd på ett matematiskt regelbundet sätt. Sannolikhetera 0,9998 och 0,86 faller då automatiskt ut