The use of databases in footwear mark cases

Outline

• What is the purpose of databases?
• Formal classification of databases
• What databases are available for issues involving footwear evidence?
• Two case-studies - interactive
• Summary Not an authoritative, exhaustive review of footwear databases
  Share ideas and explore issues

Royal Statistical Society
Errol Street, London
15th March 2016

Graham Jackson
Advance Forensic Science
and Abertay University, Dundee

g.jackson@abertay.ac.uk
advanceforensicscience@gmail.com

Adam Baines
Forensic Specialist
Lancashire Constabulary

Adam.Baines@lancashire.pnn.police.uk

© Graham Jackson, Adam Baines 2016
Footwear databases v1.8

Page 1
What is the purpose of databases?

- Provide information to assist the expert
- Provide investigative leads
- Classify material
- Answer questions of interest

- Help the expert to arrive at an opinion in relation to an issue
- Help the expert to assign probabilities

What are the types of issue that face experts?

**European Network of Forensic Science Institutes (ENFSI)**


**Association of Forensic Science Providers (AFSP)**

Standard for the formulation of evaluative forensic science expert opinion

The AFSP Standard and the ENFSI Guideline focus on the provision of ‘evaluative’ opinion...

... through the evaluation of a likelihood ratio, preferably for ‘activity’ level issues

Both documents recognise that forensic scientists also contribute to, and offer opinions relating to, investigative issues...

... but what type of opinions could, and should, scientists offer in investigative situations?

The AFSP Standard and ENFSI Guideline embrace two key concepts that formed the basis of an earlier model – ‘Case Assessment and Interpretation’ (C.A.I.)

Classification of opinion: ‘Investigative/Evaluative’ ‘Hierarchy of issues’

Based on a probabilistic, Bayesian paradigm

Types of issue

‘Investigative’

‘Evaluative’

Types of opinion

Explanations

Likelihood ratios

Posterior probabilities

There are benefits and limitations to each of these types of opinion

Categorical

Types of opinion

Explanations

Likelihood ratios

Posterior probabilities

Formal notation for the types provides guidance on what data and knowledge may help the expert

Categorical
Types of issue

Explanations: \([H_1|E,I], [H_2|E,I], \ldots, [H_n|E,I]\)

Posterior probabilities: \(\Pr[H_1|E,I], \Pr[H_2|E,I], \ldots, \Pr[H_n|E,I]\)

Categorical: \(\Pr[H_1|E,I] = 1, \Pr[H_1|E,I] = 0\)

Types of opinion

Types of issue

Evaluative

Types of opinion

\(\frac{\Pr[E|H_P,I]}{\Pr[E|H_D,I]}\)

Likelihood ratios
Footwear databases v1.8

‘Hierarchy of issues’

Did he commit the **offence**?

Did he do the **activity**?

Is he/his item the **source** of the trace material?

Is he/she the (sub)-**source** of the test result?

© Graham Jackson, Adam Baines 2016

Evaluative

... in a judicial context

© Graham Jackson, Adam Baines 2016

Most footwear issues relate to whether the suspect’s shoe left the mark.

Occasionally, the issue relates to **activity**, e.g. kicking vs treading
What is the source of the trace material?

What activity took place?

What offence was committed?

Who committed the offence?

Which other offences has offender committed?

What is the (sub)-source of the test result?

'Hierarchy of issues'

Most issues in footwear relate to what type of footwear could have left this mark?
What databases and information are available?

Who stores data/information about footwear?

- Manufacturers
- Shops/Distributors
- SATRA
- Police Forces
- National Footwear Reference Collection (NFRC)
- National Footwear Database (NFD)
- Forensic Providers
What data/information do they collect?

Manufacturers

- Pattern designs
- Sizes
- Moulds used
- Numbers produced

Source of the most comprehensive and definitive information about any particular shoe design. It could be used for frequency statements such as:

“The suspects shoe is a Nike Air Max ‘95, UK size 8.5 and has been produced in mould ‘C’. 500,000 soles have been produced in this mould. This particular style is available in 61 different colour variations.”

What data/information do they collect?

Shops/Distributors

- Distribution figures
- Sales figures

Source of frequency data such as:

200,000 of this particular type of trainer were distributed around the UK.

Taking into account the total number of shoes distributed around the UK, the relative frequency of the trainer could be assessed as 1/10 - a gross simplification.
What data/information do they collect?

**SATRA**

- Size information
  - Male
  - Female

The data are presented in percentages, e.g. 30% of women are size 6. The statistics are derived from a sample of approx. 900 women.

The data are based on foot measurements and not shoe size. This is because people often wear an incorrect shoe size for their feet. SATRA have said that women tend to wear shoes 1 size too small for them and men tend to wear shoes 1 size too big.

---

What data/information do they collect?

**Police Forces/Forensic Providers**

**Suspect database**

- Names/DOB
- Date of arrest
- Crime
- Shoe sole pattern
- Shoe size
- Gender Specific data
- Frequency
  - Pattern
  - Size
What data/information do they collect?

**Police Forces/Forensic Providers**

**Suspect database**

The data are derived from the shoes of people suspected of crime. In the case of databases held by Police Forces, the data would be specific to the Force’s geographic area and, for some Forces, could even be broken down into specific Divisional areas. Databases compiled by Forensic Science Providers would be populated by the footwear from all of the Forces submitting work to the FSP. Therefore it would cover a larger area, possibly national.

---

**Crime Scene database**

- Address
- Date of Offence
- Crime
- Sole patterns identified
- Frequency of sole patterns
What data/information do they collect?

Police Forces/Forensic Providers

Crime Scene database

The data are derived from footwear marks identified and collected at crime scenes. The data do not include footwear marks that could not be identified or those that weren't recovered. The database does not take into consideration whether the identified marks are marks made by offenders’ footwear.

What data/information do they collect?

Police Forces/Forensic Providers

Reference Collection

• Sole pattern
• Size
• Wear
• Acquired features
What data/information do they collect?

**Police Forces/Forensic Providers**

**Reference Collection**

The reference collection is often used to inform the significance of a comparison. For example, if a suspect’s shoe appears to have an unusual pattern of wear, the Reference Collection could be used to assess how many other shoes of the same sole pattern and size also show this particular wear pattern.

---

What data/information do they collect?

**National Footwear Reference Collection (NFRC)**

**Database**

- Shoe sole patterns
- Frequencies
  - Crime data
  - Custody data
What data/information do they collect?

National Footwear Database (NFD)

- Suspect data
  - Names/DOB
  - Date of arrest
  - Crime
  - Sole pattern
  - Shoe size
- Crime Scene data
  - Address
  - Date of Offence
  - Crime
  - Sole patterns identified

Investigative Case Study

- Police called to the home of an elderly woman
- A neighbour had found her body in the front living room
- The woman had been physically, sexually assaulted
- Point of entry appeared to be via a forced, ground-floor, transom window to the rear of the property
- Crime scene examiners observed a fresh-looking footwear mark in the flower bed immediately below the forced window
- Photographs and a cast of the mark were taken
What is the issue(s)?

- What type of shoe left the mark?
- What type of person left the mark?

What type of opinion would be appropriate?

Explanations?

Posterior probabilities?

Categorical?

---

Opinions after lab examination

- Shoe had a plain, or very finely-patterned, broad sole
- Most likely size was UK size 7
- Most probable style would be a suede boot
- Likely worn by an ‘older’ person

What prior hypotheses?

What prior probabilities?

What probabilities for the observations?

What data and knowledge?
• A stolen car was used to transport three offenders away from the scene of a robbery
• The car is abandoned about 1 mile away from the scene
• Crime Scene Examiners observe a footwear mark in soil on the brake pedal of the car. They believe it to have been left by the last person to have driven the car
• The owner of the car will say that he did not leave the mark
• A gel lift of the mark is taken

• A suspect was arrested about three hours after the incident
• The suspect’s footwear was taken
• The suspect denied all involvement with the crime and denied being the driver of the car
Is the suspect’s shoe the source of the mark?

H_p - The suspect’s shoe is the source of the mark
H_0 - Some other, unknown shoe is the source of the mark

Evaluation of the observation of a ‘match’ (E) at ‘source’ level

\[ LR = \frac{\Pr[E|H_p, I]}{\Pr[E|H_0, I]} \]  

Assumption of numerator approaching 1

\[ LR \rightarrow 1 \]  

A relative frequency of what?
Is the suspect’s shoe the source of the mark?

H_p - The suspect’s shoe is the source of the mark
H_d - Some other, unknown shoe is the source of the mark

\[
LR = \frac{\Pr [E | H_p, I]}{\Pr [E | H_d, I]} = \frac{1}{\text{Rel. frequency}}
\]

What sort of data would give a useful relative frequency and thereby inform the denominator probability?

What would be a relevant population in this case?


\[
LR = \frac{\Pr [\text{Match}' | H_p, I]}{\Pr [\text{Match}' | H_d, I]} = \frac{1}{\text{Rel. frequency}}
\]

Informed by data on scene marks

Informed by data on offenders like the defendant

Informed by data on innocent people like the defendant
Mark on pedal

Suspect’s shoe.

NFRC Pattern code: Ugg 17

<table>
<thead>
<tr>
<th>Population</th>
<th>Total</th>
<th>UGG 17</th>
<th>%*</th>
<th>LR**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspect</td>
<td>12089</td>
<td>23</td>
<td>0.2</td>
<td>500</td>
</tr>
<tr>
<td>Scene</td>
<td>4552</td>
<td>21</td>
<td>0.5</td>
<td>200</td>
</tr>
</tbody>
</table>

* Approximate figures
** Assuming Numerator = 1

Witnesses report the driver who ran away from the abandoned vehicle was a “small, young woman”.

The suspect is 18 years old, 5’4”.

$H_p$ - The suspect’s shoe is the source of the mark

$H_0$ - Some other, unknown shoe, worn by a “small, young woman”, is the source of the mark
### Mark on pedal

Suspect’s shoe.

NFRC Pattern code: Ugg 17

<table>
<thead>
<tr>
<th>Population</th>
<th>Total</th>
<th>UGG 17</th>
<th>%*</th>
<th>LR**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspect</td>
<td>12089</td>
<td>23</td>
<td>0.2</td>
<td>500</td>
</tr>
<tr>
<td>Scene</td>
<td>4552</td>
<td>21</td>
<td>0.5</td>
<td>200</td>
</tr>
</tbody>
</table>

* Approximate figures ** Assuming Numerator =1

### Suspect’s shoe.

NFRC Pattern code: Adidas 509

<table>
<thead>
<tr>
<th>Population</th>
<th>Total</th>
<th>Adidas 509</th>
<th>%*</th>
<th>LR**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspect</td>
<td>12089</td>
<td>1343</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Women</td>
<td>1097</td>
<td>19</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>Scene</td>
<td>4552</td>
<td>847</td>
<td>19</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Population</th>
<th>Total</th>
<th>UGG 17</th>
<th>%*</th>
<th>LR**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspect</td>
<td>12089</td>
<td>23</td>
<td>0.2</td>
<td>500</td>
</tr>
<tr>
<td>Women</td>
<td>1097</td>
<td>23</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>Scene</td>
<td>4552</td>
<td>21</td>
<td>0.5</td>
<td>200</td>
</tr>
</tbody>
</table>

* Approximate figures ** Assuming Numerator =1
Summary

Share ideas and explore issues

Not an authoritative, exhaustive review of footwear databases

- What is the purpose of databases?
  - Provide information on prior hypotheses
  - Provide data to help assign prior probabilities for hypotheses
  - Provide data to help assign probabilities for observations

- Formal classification of databases
  - ‘Investigative’
  - ‘Evaluative’

- What databases are available for issues involving footwear evidence?
  - Commercial
  - Police Forces
  - Forensic Science Providers
  - National

- Two case-studies
  - Importance of specifying the issue and the propositions
  - Impact on LR

© Graham Jackson, Adam Baines 2016
Footwear databases v1.8
41
Thanks for attention and contributions

Hope it’s been of interest to you

g.jackson@abertay.ac.uk
advanceforensicscience@gmail.com

Adam.Baines@lancashire.pnn.police.uk