Forensic Science for Biometricians

Geoffrey Stewart Morrison

Aston University
• This presentation provides a brief introduction to evaluation of strength of forensic evidence and validation of forensic analysis systems.

• The presentation assumes and is tailored for an audience familiar with signal-processing biometrics such as finger, face, and speaker recognition.
Biometrics

• Modes:
  – automatic fingerprint identification
  – automatic face recognition
  – automatic speaker recognition

• Applications:
  – security      yes / no?
  – intelligence  threshold exceeded?
  – investigative top $n$ from database searched
  – forensic evaluation strength of evidence
Biometrics

• Other applications:
  – score
  – threshold
  – binary decision

• Forensic evaluation:
  – score
  – forensic scientist does not apply a threshold or make a decision
  – likelihood ratio answering a specific question of interest to the court
Forensic evaluation

• Likelihood ratio framework:
  – Association of Forensic Science Provider in the United Kingdom and Republic of Ireland 2009 Standards for the formulation of evaluative forensic science expert opinion
  – European Network of Forensic Science Institutes 2015 Guideline for evaluative reporting in forensic science
  – 2016 report by President Obama’s Council of Advisors on Science & Technology
Forensic evaluation

**posterior odds**

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

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

**likelihood ratio**

**prior odds**
Forensic evaluation

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

\text{likelihood ratio}

\text{prior odds}

\text{responsibility of}

\text{forensic scientist}

\text{responsibility of}

\text{trier of fact}
Forensic evaluation

- Likelihood ratio value will only be meaningful if the trier of fact understands the question that it answers, and that question is a question of interest for the court.

\[
\frac{p(E \mid H_{\text{same-origin}})}{p(E \mid H_{\text{different-origin}})}
\]

Probability of acoustic properties of questioned-speaker recording if it were spoken by the known speaker versus if it were spoken by some other speaker selected at random from the relevant population.
Relevant population

LR = 1.8
Relevant population

LR = 989
Forensic evaluation

• The forensic scientist must carefully explain to the court what question the forensic scientist has set out to answer

  – the court can decide if it is a relevant question

  – the trier of fact can understand the answer
Forensic evaluation

- The forensic scientist must carefully explain to the court what data the forensic scientist has used to calculate the denominator of the likelihood ratio
  - the court can decide whether the data are sufficiently representative of the relevant population
Aside

- The statistical models use for calculating the likelihood ratio should be generative models that actually calculate the ratio of two likelihoods or discriminative models that are analogous with such generative models

  - Gaussian mixture models

  - logistic regression (analogous with linear discriminant analysis)
Validation

• A rigorous application of US Federal Rules of Evidence 702 and the criteria established by the Supreme Court in Daubert, Joiner, and Kumho Tire

• A rigorous application of England & Wales Criminal Practice Directions 19A

– would require empirical validation under conditions reflecting those of the case under investigation
Validation

- See also:
  - 2016 report by President Obama’s Council of Advisors on Science & Technology
  - Forensic Science Regulator of England & Wales 2014 Guidance on validation
  - European Network of Forensic Science Institutes 2015 Methodological guidelines for best practice in forensic semiautomatic and automatic speaker recognition
Validation

- The forensic scientist must carefully explain to the court what data the forensic scientist has used to test the performance of the forensic analysis system

  - the court can decide whether the data are sufficiently reflective of the conditions of the case that the test results will be meaningful indicators of the performance of the system under the conditions of the case

  - the court can decide whether the performance of the system under those conditions is good enough
Aside

• The training data should also reflect the conditions of the case

  – in other applications the system has to be generalizable

  – in a forensic application only one trial matters: the comparison of the actual known-origin sample and questioned-origin specimen
Validation

- Correct-classification / classification-error rate is not appropriate
  - based on posterior probabilities
  - hard threshold rather than gradient

<table>
<thead>
<tr>
<th>fact</th>
<th>decision</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>same</td>
<td>correct</td>
<td>false</td>
</tr>
<tr>
<td></td>
<td>acceptance</td>
<td>rejection</td>
</tr>
<tr>
<td>different</td>
<td>false</td>
<td>correct</td>
</tr>
<tr>
<td></td>
<td>acceptance</td>
<td>rejection</td>
</tr>
</tbody>
</table>
Validation

![Graph showing classification error rate versus Log$_{10}$ Posterior Odds](image)

- **Classification error rate** is plotted on the y-axis.
- **Log$_{10}$ Posterior Odds** is plotted on the x-axis.
- The graph distinguishes between **miss** and **false alarm** outcomes.

The y-axis indicates the classification error rate, while the x-axis represents the logarithmic scale of the posterior odds, ranging from -3 to 3.
Validation

$C_{llr}$

Log$_{10}$ Likelihood Ratio

Validation
Conclusion

• No binary decisions

• Likelihood ratio answering a specific question

• Question asked and data used to answer question must be communicated to court

• Empirical validation using data that reflect the conditions of the case

• Validation metrics consistent with the likelihood ratio framework
Thank You

http://geoff-morrison.net/