Panel on the Use of Interval Quantifications for the Value of Forensic Evidence

Danica Ommen

Ph.D. Candidate
South Dakota State University

May 6, 2016

Disclaimer

The research detailed in this presentation was supported in part by the National Institute of Justice, Office of Justice Programs, US Department of Justice under Award No. 2014-IJ-CX-K088.

The opinions and conclusions or recommendations expressed in this presentation are those of the author and do not necessarily represent those of the Department of Justice.

My Philosophy

- Bayesian-Frequentist Fusion
- Decision Theory

Questions

1. "What is the parameter we are constructing an interval for when we present an interval for the value of evidence?"
   OR
   "What additional information does an interval capture about the value of evidence?"
2. "How does a decision maker use an interval to make a decision in a logical and coherent manner?"
3. "Does presenting an interval quantification of the value of forensic evidence cause any harm?"

Question 1 - Part 1

1. **"What is the parameter we are constructing an interval for when we present an interval for the value of evidence?"**

   OR

   "What additional information does an interval capture about the value of evidence?"

2. "How does a decision maker use an interval to make a decision in a logical and coherent manner?"

3. "Does presenting an interval quantification of the value of forensic evidence cause any harm?"

Question 1 Response

- The Bayes Factor

$$V = \frac{\pi(e|H_p)}{\pi(e|H_d)} = \frac{\int f(e|\theta, H_p)\pi(\theta|H_p)d\theta}{\int f(e|\theta, H_d)\pi(\theta|H_d)d\theta}$$

- The Likelihood Ratio

$$\lambda(\theta_0) = \frac{f(e_u|\theta_{p_0})}{f(e_u|\theta_{d_0})}$$

Question 1 - Part 2

1. "What is the parameter we are constructing an interval for when we present an interval for the value of evidence?"

   OR

   "What additional information does an interval capture about the value of evidence?"

2. "How does a decision maker use an interval to make a decision in a logical and coherent manner?"

3. "Does presenting an interval quantification of the value of forensic evidence cause any harm?"

Question 1 Response

- Use the interval as a computational technique to obtain a reliable numerical solution
- Example: Monte Carlo Standard Error for the BF

  Let $\hat{V}$ be a numerical approximation via MC integration of V.
  Let ϵ_V be the MCSE of $\hat{V}$ for V.
  Present $\hat{V}$ as reliable numerical solution to V when

$$[\hat{V} - 2\epsilon_V, \hat{V} + 2\epsilon_V]$$

  is sufficiently "short".

Question 2

1. "What is the parameter we are constructing an interval for when we present an interval for the value of evidence?"

   OR

   "What additional information does an interval capture about the value of evidence?"

2. **"How does a decision maker use an interval to make a decision in a logical and coherent manner?"**

3. "Does presenting an interval quantification of the value of forensic evidence cause any harm?"

Question 2 Response

- Use the most conservative endpoint of the interval
- Example: Single suspect vs. Single alternative source

H_p: The trace came from Fred.
H_d: The trace came from Bob.

$$\pi(H_p) = 1 - \pi(H_d) = 0.9$$

$$5 \leq \lambda(\theta) \leq 100$$

$$\begin{aligned} \frac{\pi(H_p|e)}{\pi(H_d|e)} &= \hat{\lambda}(\theta) \times \frac{\pi(H_p)}{\pi(H_d)} \\ &= 5 \times \frac{9}{1} \\ &= 45 \end{aligned}$$

Probability Fred did it ≈ 0.98

H_p: The trace came from Bob.
H_d: The trace came from Fred.

$$\pi(H_p) = 1 - \pi(H_d) = 0.1$$

$$1/100 \leq \lambda(\theta) \leq 1/5$$

$$\begin{aligned} \frac{\pi(H_d|e)}{\pi(H_p|e)} &= \frac{1}{\hat{\lambda}(\theta)} \times \frac{\pi(H_d)}{\pi(H_p)} \\ &= 100 \times \frac{1}{9} \\ &= 100/9 \approx 11 \end{aligned}$$

Probability Fred did it ≈ 0.92

Question 2 Response

- Use the midpoint of the interval
- Example: Blood-type

  Consider a single sample bloodstain trace has been recovered.
  Let θ be probability that random person's profile matches trace.
  $n = 345$ is number people whose profiles were sampled.
  $x = 29$ is number people observed to have profile matching trace.
  Let $\lambda(\theta) = 1/\theta$ be the LR.

$$x|\theta \sim Bin(n, \theta),\ \ \theta \sim Beta(0.5, 0.5) \implies \theta|x \sim Beta(x{+}0.5, n{-}x{+}0.5)$$

Table : Credible Intervals for the Likelihood Ratio in Blood Type Example

Method	Center	Lower	Upper	Width
HPD	12.36669	8.200185	16.533202	8.333017
ET	12.84116	8.552564	17.129762	8.577198
Norm	12.09746	7.774548	16.420377	8.645829

Question 3

1. "What is the parameter we are constructing an interval for when we present an interval for the value of evidence?"

   OR

   "What additional information does an interval capture about the value of evidence?"

2. "How does a decision maker use an interval to make a decision in a logical and coherent manner?"

3. **"Does presenting an interval quantification of the value of forensic evidence cause any harm?"**

Question 3 Response

- Let $\lambda_{ss}(\theta)$ be the LR as a function of θ and V_{ss} be the BF

$$\begin{aligned}
E(\lambda_{ss}(\theta)|e_s,\, e_a) &= \int \lambda_{ss}(\theta) d\Pi(\theta|e_s, e_a) \\
&= \int\int \frac{f(e_u|\theta_s)}{f(e_u|\theta_a)} d\Pi(\theta_a|e_a) d\Pi(\theta_s|e_s) \\
&= \int f(e_u|\theta_s) d\Pi(\theta_s|e_s) \int \frac{1}{f(e_u|\theta_a)} d\Pi(\theta_a|e_a) \\
&\geq \int f(e_u|\theta_s) d\Pi(\theta_s|e_s) \frac{1}{\int f(e_u|\theta_a) d\Pi(\theta_a|e_a)} \\
&= V_{ss}
\end{aligned}$$

- The "midpoint of the interval" for the LR is overstating the value of evidence
- The interval quantification is biased against the suspect

Recommendations

1. Do not present intervals in court as a surrogate for the Bayes Factor
   - Intervals themselves cannot be used to make a reasonable decision
   - Making a decision based on a credible interval for the Likelihood Ratio will be biased against the suspect
   - Only use intervals as a computational technique to get a reliable numerical answer
2. Present the Bayes Factor and the estimate of numerical precision
   - If you can't get the Bayes Factor, try the Neyman-Pearson Likelihood Ratio
   - Be honest/upfront about your methods if you present an adhoc solution in court
3. Make decisions based on the Bayes Factor
   - There is a well-defined statistical framework around it
   - Decisions based on it will be reasonable