STATISTICS EMPOWERS CRIMINAL JUSTICE

STATISTICAL SIGNIFICANCE

Federico Veneri
Ph.D. Student in Statistics
Research assistant
CSAFE-ISU

Danica Ommen
Assistant professor
Principal investigator
CSAFE-ISU

A crime was committed, and forensic evidence was found at the scene. Cue the sunglasses, The Who soundtrack, and the statisticians.

THE FORENSIC QUESTION

In a criminal case, forensic scientists examine the evidence and present findings to jurors, who will make the final judgment regarding the guilt or innocence of an individual.

The research is often related to the source of the evidence:
-Did these bullets come from the suspect's gun?
-Did the same person write these two threatening notes?

Forensic experts will try to assess similarities between the two items and provide a concluding statement. Is there sufficient evidence to say whether they come from the same source or not?

THE (LACK OF) STATISTICS IN FORENSICS

Establishing similarities between two items is a complex task. In some forensic domains, comparisons are made based on visual inspection that relies on years of experience and training.

Studies show that: experts don't always agree, some domains don't have a robust scientific foundation, and the misuse of forensic science has played a role in wrongful convictions.

As a result, several calls have been made to strengthen the scientific basis and statistical foundations in criminal justice.

HERE COMES THE STATS

Experts recommend that forensic findings be based on sound statistical foundations and a probabilistic framework to communicate the uncertainty behind the conclusions.

Likelihood ratios are advocated to measure the probative value of the evidence, but formulating a statistical model is not always feasible, specifically for complex evidence.

Statisticians have contributed new tools that assess the similarity between items based on machine learning algorithms and computer vision.

These similarity scores are the basis for computing Score Likelihood Ratios. They allow us to assess the evidence's probative value when we need an alternative to a probabilistic model.

BUT ARE THE ASSUMPTIONS MET?

The standard approach to developing Score Likelihood Ratios first creates a data set of all possible pairwise comparisons from an available background population sample.

However, using all comparisons results in complex dependence structures.

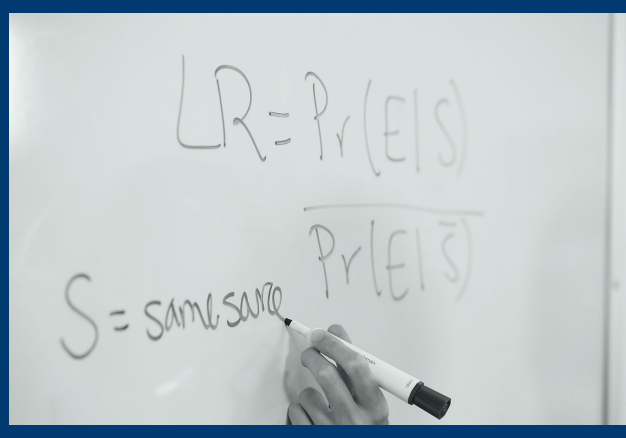
The independence assumption required by popular machine learning algorithms and density estimation procedures is no longer met.

A SAMPLING REMEDY

Our work introduces a sampling algorithm that remediates the dependence structure in forensic comparisons.

Like ensemble learning, multiple base algorithms learn over a partial view of the data (where assumptions are met), and scores are later aggregated into a final conclusion.

Ensemble Score Likelihood Ratios perform better than their traditional counterparts: providing stronger, more stable, and less misleading evidence.



BUILDING FOUNDATIONS

Statistics play a crucial role in forensic science and criminal justice. Misuses can wrongfully convict an innocent person or let a guilty person walk.

Statisticians can significantly contribute to strengthening the foundations of a more transparent, reliable, and fair criminal system.

We expect to extend our work to other forensic domains and statistical problems that depend on pairwise comparisons.

Want to learn more ?

Read previous research by CSAFE

Learn about OSAC at NIST

Read a review of issues in forensic science by the Innocence Project