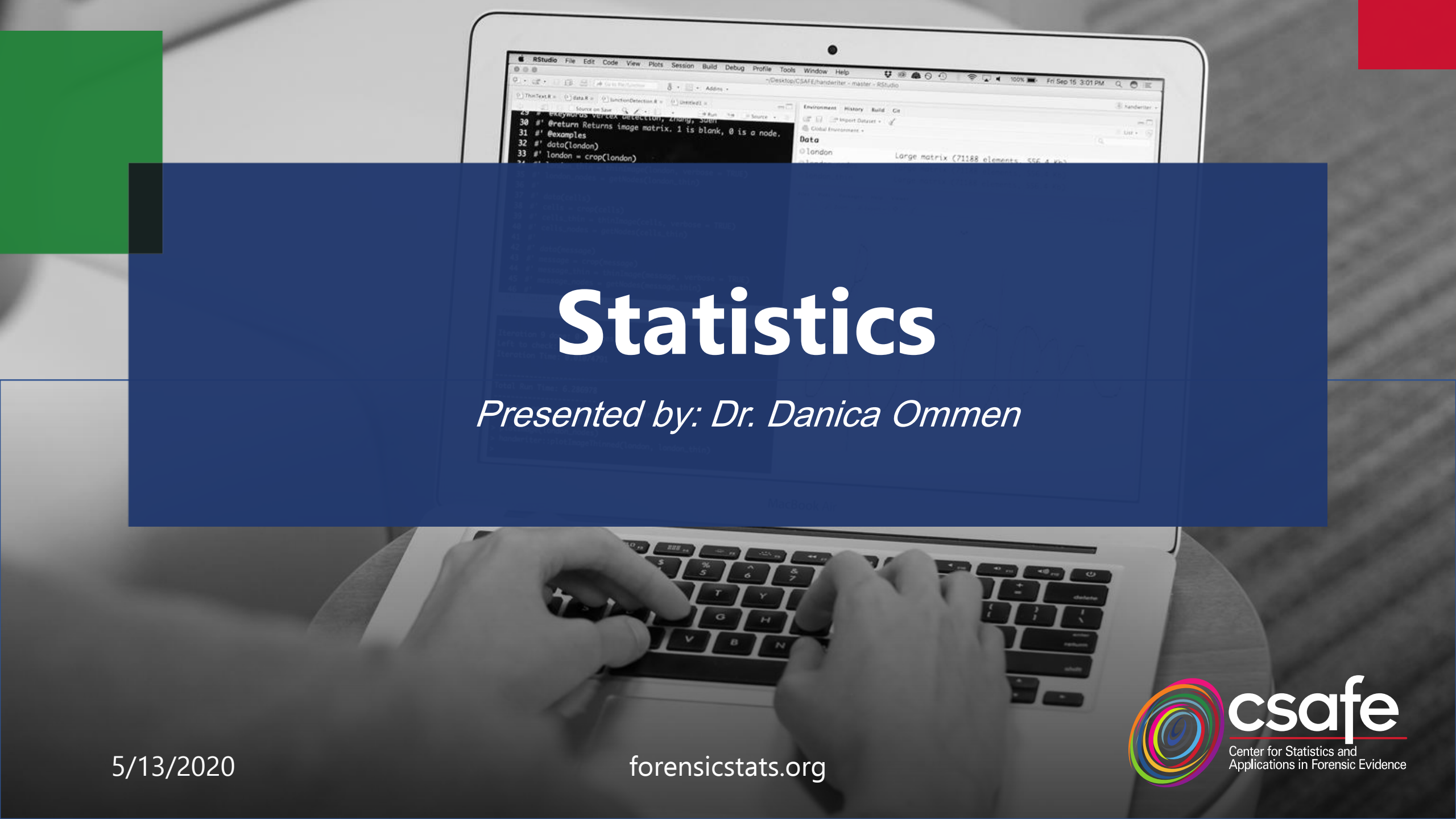

Statistics

Presented by: Dr. Danica Ommen

5/13/2020

forensicstats.org

csafe
Center for Statistics and Applications in Forensic Evidence

What is the “Statistics” group discussion about?

- Focus is on topics that are:
  - Applicable to multiple disciplines
  - Would benefit from more in-depth statistical investigation

- The research objectives that we describe will
  - Solve problems faced by a number of the other CSAFE project areas
  - Create collaborative opportunities with other projects

CSAFE 1.0 Accomplishments

Primary Focus: forensic databases

- Databases can provide analysts with information about the variability of measurements, rarity of characteristics, etc.
- Databases are necessary for proper assessment of evidential value
- There are challenges:
  - Are current databases, often collected primarily via convenience samples, well-suited to this type of analysis?
  - Are current databases easily accessible to researchers and practitioners?

Other projects (publications / presentations):

- Method of reporting conclusions regarding value of evidence using a "pool-reduction strategy"
  - At each step in the process, how many sources can be eliminated?
- Statistical approach (using probabilistic graphical structures) for understanding complex evidence

CSAFE 1.0 Accomplishments

Major Accomplishments:

- Convened an international meeting on issues related to assessing forensic databases and their use in casework, held at CMU in Pittsburgh.
- Framework for evaluating the significance of a forensic comparison published in Significance, highlighting:
  - What data are needed for evaluation of a match with respect to a population of interest.
  - The need for a database to assess the probability of seeing similar results by chance, either to develop a random match probability or estimate the denominator in a likelihood ratio.
  - Lack of availability of relevant databases in many pattern disciplines
- PhD dissertation (Tackett, UVA) on Bayesian statistical methods for analyzing fingerprint database searches

Impact:

- Deepened understanding of need for publicly available databases.

CSAFE 2.0 Statistics Projects and Lead Investigators

STAT I- Statistical Methods to Assess Reliability, Reproducibility, Accuracy of Categorical Forensic Opinions

- Lead PI: Hal Stern, UCI

STAT II- Validation and Reliability of Score-Based Likelihood Ratios (SLRs) for Forensic Evidence

- Lead PI: Danica Ommen, ISU

STAT III- Machine Learning Methods for Dependent Score-Data Resulting From Forensic Evidence Comparisons

- Lead PI: Danica Ommen, ISU

Research Area Objectives

GOAL:
Critically examine statistical issues related to methods for analyzing pattern and digital evidence that are relevant across CSAFE project areas

- Firearms
- Footwear
- Handwriting
- Digital

(Note: The issues considered here are motivated primarily from statistical tools developed during the initial CSAFE funding period.)

Projects focus on issues at three different "levels" of the forensic evidence evaluation process

- Framework for evidence interpretation (score-based likelihood ratios)
- Methods of assessing (dis)similarity (inference for machine learning techniques)
- Expressing and assessing conclusions (reliability / validity of categorical conclusions)

Background

- **Likelihood ratio (LR)** as a summary of the evidence:

$$LR = Pr(E \mid H_s) / Pr(E \mid H_d)$$

  - The numerator assesses how likely is the evidence if H_s is true
  - The denominator assesses how likely is the evidence if H_d is true
- Likelihood ratios are challenging to apply with pattern / digital evidence
  - Mathematical representation of the evidence (often an image) is very high-dimensional
  - Developing probability models for E is challenging
- **Score-based likelihood ratios (SLR)** have been proposed by some
  - Replace the evidence E by a "score" S summarizing differences/similarities of the two samples
  - $SLR = Pr(S \mid H_s) / (Pr(S \mid H_d)$
- Several CSAFE 1.0 projects developed a score by using **machine learning algorithms** to distinguish between known match pairs and known non-match pairs; the estimated probability of being a match is a possible score

CSAFE 2.0

Motivation - Framework for Evidence Interpretation:

- Likelihood ratio or Bayes factor have been proposed as the ideal approach to analyzing forensic evidence
- Not generally possible to develop full probability models for pattern and digital evidence
- Much attention focused on score-based likelihood ratios, but statistical properties not well understood

Proposed Activities:

- Explore the strengths and weaknesses of SLRs for quantifying the value of evidence
- Develop framework of evidence interpretation which exploits the strengths of SLRs for pattern and digital evidence

Potential Impact:

- List of recognized strengths and weaknesses, with supporting statistical arguments, of SLRs
- Framework for expressing conclusions regarding SLR results
- Applicable to a wide range of impression and pattern evidences (e.g., CSAFE 1.0 projects in footwear and firearms)

CSAFE 2.0

Motivation - Methods for Assessing Similarity (Dissimilarity) of Evidence Samples

- Machine learning (ML) methods are useful tools for assessing similarity of pattern evidence in various disciplines
- ML methods are fit to training sample comprised of pairs of known matches and known non-matches
- Standard algorithms rely on the assumption of independence of the cases used to fit the models to the data
- Assumption does not hold when we construct all possible pairs (mated and non-mated) of a set of items

Proposed Activities:

- Explore the extent to which violating the independence assumption affects the performance of ML methods
- Develop ML methods for evaluating comparison scores that accommodate/adjust for the dependency in the data

Potential Impact:

- Provide statistically rigorous methods of computing SLRs for variety of evidence types
- Critically evaluate CSAFE methods for potential areas of correction/improvement before deploying methods in labs

CSAFE 2.0

Motivation - Expressing and Assessing Conclusions:

- In the absence of convincing likelihood ratio methods, pattern evidence disciplines rely on categorical conclusion scales

Proposed Activities:

- Develop statistical approaches to the analysis of data from reliability (repeatability and reproducibility) studies for categorical conclusions in forensic science.
- Develop statistical approaches to the analysis of validity studies for categorical conclusions in forensic science.
- Apply the statistical approaches to available data from the pattern disciplines. Data can include publicly available data from the FBI fingerprint study (and perhaps additional data from new studies of shoeprints, handwriting).

Potential Impact:

- Most studies of methods have focused on outcomes of binary decisions
- Statistical approaches to assessing reliability and validity for categorical scales will be important in understanding their properties.

Resources and Needs

Impact Summary:

- Our results are applicable to a wide range of evidence types
- Statistically sound methods of expressing evidential value
- Better methods for communicating results
- Support for decisions/conclusions

Resources and Needs

Ways forensic science community can engage/partner with us:

- Engage/partner in discipline specific projects
  - Firearms
  - Footwear
  - Handwriting
  - Digital
- Data
  - Consider collaborating with us on analysis of existing data that you may have
    - Case studies
    - Black box studies
    - Proficiency tests
  - Participate in CSAFE data collections