B90 An Evaluation of Score-Based Likelihood Ratios (SLRs) for Glass Data

Federico Veneri, MSc, CSAFE - Iowa State University, Ames, IA 50011; Danica Ommen, PhD, Iowa State University - Statistics, Ames, IA 50011*

Learning Overview: The goal of this presentation is to illustrate the use of statistical simulation to study the behavior of SLRs. Using realistically simulated data of trace elemental analysis for glass, this presentation compares the behavior of different SLRs to their likelihood ratio counterpart under the common source scenario and assesses the dependence structure that may arise due to the choice of a training sample.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by providing attendees with a review of tools for assessing SLRs and dependence structure that may arise in their use.

The use of likelihood ratios is advocated as a way to provide a numeric assessment of the evidential strength during the forensic expert testimony. The development of such ratios requires the construction of probability models that might be challenging, if not infeasible, to estimate. To avoid this, researchers can implement similarity scores as a way of reducing the complexity of a model into a potentially lower-dimensional metric, but there is still concern regarding behavior of such a score-based likelihood ratio.[1,2]

Consider the following scenario where a forensic examiner evaluates two glass fragments. We can frame this problem under the common source problem as follows: given two fragments (X and Y) that come from their respective windows (W1 and W2), examiners will deal with the hypothesis that windows one and two are the same, meaning that the fragments come from the same source, or the alternative hypothesis that they are different.

Without relying on normality assumptions, the researcher may approach the problem by computing a similarity score, a quantitative measure of how similar the chemical compositions of two fragments are. Previous projects have proposed using a random forest similarity score, while others that study chemical composition have used distance-based scores when addressing chemical composition.[3-5] This score gives the examiner a quantitative input. Still, a reference is required to know if the score found is more likely under the same source or the different source hypothesis.

The approach described requires two key components: (1) the development of a similarity score, and (2) estimating the distribution of such scores under the same source or different source scenarios. For the second component, the researcher can construct a database of pairwise comparison where the ground truth is known, meaning if the pair is a known match (same source) or a known non-match (different source).

Inquiry regarding the behavior of SLR has been previously made in the literature.[2,6] A factor less explored is the dependence on the training database. To assess this, this study created realistically simulated data using the chemical composition of glass data.[7,8]

Using an approach inspired by machine learning, three datasets were simulated that play the role of training, testing, and validation sets to compare four score-based likelihood ratios: (1) Euclidean distance, (2) city distance; (3) random forest score, and (4) a random forest with an additional down sampling step during the training stage.[9]

In each simulation, this study used its designated training set to train the random forest model, the test set to estimate the distributions under both hypotheses. Lastly, this study used the validation set to illustrate two critical aspects. First, since the data was generated following a known model, this study compared the decision reached using the SLR and the corresponding LR in terms of the rate of misleading evidence and if they fall in the same range of strength of evidence. Secondly, using permutation of the simulated training, testing, and validation sets, this study presents measures of dependence on the training data.

Reference(s):

1. Hal S Stern. Statistical Issues in Forensic Science. *Annual Review of Statistics and Its Application* 4 (2017): 225–244.
2. Cedric Neumann and Madeline Ausdemore. Defence against the Modern Arts: The Curse of Statistics—Part II: 'Score-Based Likelihood Ratios. *Law, Probability and Risk* 19, no. 1 (March 1, 2020): 21–42, https://doi.org/10.1093/lpr/mgaa006.
3. Soyoung Park and Sam Tyner. Evaluation and Comparison of Methods for Forensic Glass Source Conclusions. *Forensic Science International* 305 (December 2019): 110003, https://doi.org/10.1016/j.forsciint.2019.110003.
4. Soyoung Park and Alicia Carriquiry. Learning Algorithms to Evaluate Forensic Glass Evidence. *The Annals of Applied Statistics* 13, no. 2 (June 2019): 1068–1102, https://doi.org/10.1214/18-AOAS1211.
5. Annabel Bolck, Haifang Ni, and Martin Lopatka. Evaluating Score- and Feature-Based Likelihood Ratio Models for Multivariate Continuous Data: Applied to Forensic MDMA Comparison. *Law, Probability and Risk* 14, no. 3 (September 1, 2015): 243–66, https://doi.org/10.1093/lpr/mgv009.
6. N. Garton, D. Ommen, J. Niemi, and A. Carriquiry. Score-based likelihood ratios to evaluate forensic pattern evidence. *arXiv preprint arXiv:2002.09470*, 2020.
7. Danica M. Ommen, Christopher P. Saunders, and Cedric Neumann. The Characterization of Monte Carlo Errors for the Quantification of the Value of Forensic Evidence. *Journal of Statistical Computation and Simulation* 87, no. 8 (May 24, 2017): 1608–43, https://doi.org/10.1080/00949655.2017.1280036.
8. C. G. G. Aitken and D. Lucy. Evaluation of Trace Evidence in the Form of Multivariate Data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 53, no. 1 (January 2004): 109–22, https://doi.org/10.1046/j.0035-9254.2003.05271.x.
9. Didier Meuwly, Daniel Ramos, and Rudolf Haraksim. A Guideline for the Validation of Likelihood Ratio Methods Used for Forensic Evidence Evaluation. *Forensic Science International* 276 (2017): 142–53, https://doi.org/10.1016/j.forsciint.2016.03.048.

Score Likelihood Ratios, Learning Algorithms, Glass Evidence

*Presenting Author