Applications of goodness-of-fit testing in forensic source identification problems

Janean Hanka
(South Dakota State University, Brookings, USA)

December 11, 2026

In forensic source identification problems, it is common to encounter large-scale few-shot learning scenarios, where we have a large number of sources with very few high-dimensional objects per source. To work in a more tractable space, a score function is used between all pairs of observations, reducing the problem to the real number line. However, there are still relatively few observations, and there is now an induced dependency structure. Due to these difficulties, it is often assumed that all within-source scores follow a common distribution. This common within-source distribution of scores is then used in forensic evidence interpretation methods, such as the classical two-stage approach, and the (score-based) likelihood ratio approach.

Recently, we've developed goodness-of-fit tests designed to test the assumption of a shared within-source distribution of scores. These tests rely on invariant statistics and parametric bootstraps, making them applicable to the small sample sizes commonly seen in forensic science. This presentation focuses on the one-sample goodness-of-fit test for model selection, and the ways to build up the distribution of within-source scores relative to a specific source. We discuss the theory behind the test, the consistency and potential super efficiency of the estimate of the nuisance parameter, and the implications of the test with respect to forensic evidence.