A159 Multivariate Empirical Bayes Methods Applied to the Identification of Impression Evidence

Brooke W. Kammrath, PhD, University of New Haven, Forensic Science Dept, 300 Boston Post Rd, West Haven, CT 06516; Frani Kammerman, MS, John Jay College of Criminal Justice, 454 West 59th St, New York, NY 10019; Loretta Kuo, BS, 899 10th Ave, New York, NY 10019; Peter J. Diaczuk, BS, 445 W 59th St, New York, NY 10019-2925; Patrick McLaughlin, BS, John Jay College of Criminal Justice, 454 West 59th St, New York, NY 10019; Elizabeth W. Phelps, MFS, Boston, MA; James E. Hamby, PhD, Int'l Forensic Science Lab, 410 Crosby Dr, Indianapolis, IN 46227; and Nicholas D. Petraco, PhD, John Jay College of Criminal Justice, 454 West 59th St, New York, NY 10019*

After attending this presentation, attendees will better understand how to apply empirical Bayes multivariate statistical methods to the identification of impression evidence, specifically firearm and tool markcomparisons.

This presentation will impact the forensic science community by presenting examples and explanations of empirical Bayes methods for the identification of impression evidence.

For several decades, and especially since the National Academy of Sciences' 2009 Report *Strengthening Forensic Science in the United States*: *A Path Forward*, forensic firearm and tool mark comparisons have been under increased scrutiny. A significant criticism of the analysis of pattern and impression evidence is that there is no accepted methodology to generate numerical proof that independently corroborates morphological conclusions. This research critically evaluates the use of empirical Bayes methods for the analysis of firearm and tool mark impression evidence.

For more than half a century, empirical Bayes methods have been applied to many types of problems in a plethora of disciplines including, but not limited to, genomics, economics, epidemiology, safety engineering, and quality assurance. Its rise in popularity can be attributed to the increase in data set sizes, computing power, and because of the straightforward nature of inference based on Bayes' rule. However, empirical Bayes methods have not yet been applied to the statistical analysis of forensic evidence.

The appeal of empirical Bayes methods is that it blends objective frequentist and subjective Bayesian approaches which lead to a falsifiable inference method consistent with the Popperian philosophy of science. Empirical Bayes methods can be interpreted as an approximation to a "fully Bayesian treatment" of a hierarchical model in which the parameters at the highest level of the hierarchy are set to their most likely values, rather than being integrated out. From a forensic science perspective, the most significant advantage of empirical Bayes methods is their ability to formulate clear statistical inferences in which the prior probability is estimated objectively from the data rather than subjectively before the data is observed.

In this research, 3D quantitative surface topographies of firearm and tool markstriated impressions were collected using confocal microscopy. A reasonably complete striation pattern was then summarized as multivariate feature vectors in the form of mean profiles. Next, principal component analysis (PCA) was used to reduce correlation within each profile while maintaining the essential information and a support vector machine (SVM) learning algorithm was used for identification. Last, empirical Bayes methods were used to estimate local false discovery rates (FDR), also known as posterior error probabilities, which quantitatively assesses whether an identification *really is correct*. In order to alleviate the concerns of data reuse, common in empirical Bayes based schemes, we adopted the standard multivariate model fitting practice of first splitting the data into separate training, validation and test sets. The training set was used to generate estimates of a null likelihood and null prior. The validation set was then used to find FDR estimates based on the training set null estimates. This approach has the added advantage that estimates of standard errors for fdrs are available. At each step in the process, goodness-of-fit diagnostics were applied and independence assumptions on the fit z-values (required for the standard error formulas) were tested. The FDR estimates can ultimately be used to accompany each identification of an unknown with an output by a machine learning algorithm, providing statistical support for impression evidence conclusions.

Empirical Bayes, Impression Evidence, Statistics