NATIONAL COMMISSION ON FORENSIC SCIENCE

Views of the Commission
Statistical Statements in Forensic Testimony

Subcommittee
Reporting and Testimony
Status
Initial Draft

Date of Current Version	19/07/16
Approved by Subcommittee	22/08/16
Approved by Commission	[*dd/mm/yy*]

Note: This document reflects the views of the National Commission on Forensic Science, and does not necessarily represent the views of the Department of Justice or the National Institute of Standards and Technology. This document does not formally recommend any action by a government entity, and thus no further action will be taken upon its approval by the Commission.

Overview

This views document presents background information on the following question: When experts present the results of forensic science examinations, tests, or measurements, what quantitative or qualitative statements of probability should they provide?

The 2009 report of the NRC Committee on Identifying the Needs of the Forensic Sciences Community emphasized the importance of describing uncertainties in measurements and inferences. The mathematical theories of probability and statistics provide a framework for reasoning and expressing conclusions about uncertain events or facts. In mathematics, probability is simply a function that associates numbers between 0 and 1 with events or outcomes and that conforms to a set of mathematical axioms. These mathematical probabilities have been used to describe the chances that particular events will occur and that statements about events or hypotheses are true. They also have been used to quantify the degrees of belief that a person should have in these matters. Statistics is concerned with the study of variability, with the study of uncertainty and with the study of decision-making in the face of uncertainty. Its focus is a set of principles for drawing scientific conclusions from data based on probability.

Forensic testimony at trial may include statistical statements regarding different points at issue. The most common of these involves trace evidence and pattern matching, but there are many others such as estimating the time of death, the statistical interpretation of the origin of a weapon from blood spatter, or causal attribution associated with an observed outcome. What all of these have in common are: 1) the existence of a relevant database describing related objects (e.g., finger prints) or observed data or experimental results, 2) a statistical model for the question at hand (e.g., comparing DNA samples from crime scenes with those of a suspect using the probability model for alleles in DNA), 3) information on measurement error associated with the evaluation of the forensic evidence, and 4) a statistical statement regarding the probative value of any comparisons done or calculations performed (e.g., how rare is an observed positive association?).

For trace or pattern evidence, the statistical model needs to address the validity and reliability of the process/method used to determine whether there is a positive association between the questioned sample (e.g., crime scene evidence) and the known sample (e.g., a reference sample from the defendant). Some refer to this as the foundational validity of the forensic method, and this is typically established via one or more scientific studies, whose details and results are vetted by an independent scientific organization[1]. For most of trace and pattern evidence today there is no commonly accepted probabilisitic model and most effort to generate probabilistic statements come from empirical studies of matching performed on the database.

All of the statistical calculations should be replicable given the data and statistical model, whereas the quantitative summary of the actual forensic evidence may vary from examiner to examiner and from laboratory to laboratory. Such measurement error should be an integral part of the expert report.

What to Report?

For DNA, and other trace and pattern evidence the primary focus of the forensic examination is on “matching” the crime scene evidence and the reference sample from the defendant. But a high probability of a “match” does not make the evidence probative unless a match is also rare in the relevant database defined for the case at hand. Thus much statistical attention has focused on the reporting of the likelihood ratio, e.g., the ratio of the match probability to the probability of a match with a randomly chosen person from the relevant population. Different choices of a relevant population or the use of different databases will yield different values of the likelihood ratio.

Although likelihood ratios are by no means confined to DNA evidence[2] and could in principle be constructed in most cases involving forensic evidence, they appear most often in DNA-mixture cases[3] and in paternity or other kinship cases (where the numerator of the *LR* is not 1).[4]

---

[1] National Commission on Forensic Science Views Document on Technical Merit Evaluation of Forensic Science Methods and Practices.Adopted at NCFS Meeting #10 - June 21, 2016. https://www.justice.gov/ncfs/file/881796/download

[2] See NIST Expert Working Group on Human Factors in Latent Print Analysis, Latent Print Examination and Human Factors: Improving the Practice Through a Systems Approach (David H. Kaye ed., 2012) (suggesting a likelihood format for presenting the findings of latent-print examiners).

[3] Some courts have held that probabilities need not be provided in such cases—the analyst may testify that the defendant could not be excluded as a contributor to a mixed DNA sample without giving a probability of exclusion or other statistic. See, e.g., Rodriguez v. State, 273 P.3d 845, 851 (Nev. 2012) (discussing cases).

[4] E.g., State v. Ott, 80 So.3d 1280, 1285 (La. Ct. App. 2012) (“the defendant was ultimately determined to be at least 14.3 million times more likely than another random person to have contributed to” the mixture); David H. Kaye,

Internationally, many forensic science laboratories and scientists favor reporting likelihood ratios. Efforts have been made to map intervals of *LR* values into phrases such as "limited evidence" (for *LR*s of 10 of less) or "extremely strong evidence" (for *LR*s of a million or more)[5] and to use the same scale for rough estimates of likelihoods made without relying on systematically collected data.[6] As with random-match probabilities, judicial opinions often mischaracterize *LR*s as statements of probabilities or odds in favor of a match. Even for DNA, once the possibility of multiple sources arises, there exists a multiplicity of proposals for how to report a LR, or a related quantity from a Bayesian analysis called a Bayes factor.

There is no simple and agreed upon statistic to report associated with trace or pattern evidence. While some statisticians strongly advocate the use if a specific form of the LR others propose variants of this quantity

Any recommendation on presenting explicit probabilities or likelihood ratios in light of forensic evidence might distinguish between probabilities based on some statistical model and ones said to flow from the forensic evidence itself. Does a forensic technician testifying in court actually compute significance probabilities or likelihood ratios, or do they come from a computer program developed by statisticians and related forensic experts? If the statistical model is known to be at best an approximation, how should the probabilistic statements coming from it be viewed? What if the statistical model and method used to analyze the evidence do not admit naturally to the simplistic form of likelihood ratio increasingly favored in the forensic-science literature? How should probabilistic statements be viewed if they are not based on all of the "relevant evidence"?

Views of the Commission

It is the view of the commission that:

1. No one form of statistical statement is most appropriate to all forms of pattern and trace evidence, and thus the expert needs to be able to support, either via a report or by direct testimony, the choice made.

2. More importantly, the forensic expert, reporting whatever statistical quantity, needs to be able to also report on the uncertainty associated with it in some form. This might take the form of a reported interval or more typically separate statements regarding errors and

---

David A. Bernstein & Jennifer L. Mnookin, The New Wigmore: A Treatise on Evidence: Expert Evidence § 14.2.2 (2d ed. 2011).

[5] See Colin Aitken & Franco Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists 107-08 (2d ed. 2004); Association of Forensic Science Providers, Standards for the Formulation of Evaluative Forensic Science Expert Opinion, 49 Science & Justice 161–164 (2009); see also Regalado v. Estate of Regalado, 933 N.E.2d 512, 523 n.7 (Ind. Ct. App. 2010) ("likelihood ratio between 10 and 100 indicates 'Moderate support' for hypothesis that Paula and Tony are half-siblings").

[6] Cf. R v. T [2010] EWCA Crim 2439 (disapproving of testimony of "moderate scientific support" based on a likelihood ratio of 100 derived from some data on the tread patter and size of shoes in England and personal estimates of the rarity of other features).

uncertainties associated with the analysis of the evidence and not simply the variations in the likelihoods themselves.

3. Forensic experts should present and describe the similarities and differences in the feature sets of the questioned and known samples (the data).

4. Forensic experts **should not** state that a specific individual or object is the source of a trace without explaining that it is possible that other individuals or objects could possess or have left a similar set of observed features.

5. To explain the value of the data in addressing hypotheses as to the source of a questioned sample, forensic examiners may:

   A. Refer to relative frequencies of features in a sample of individuals or objects in a relevant population (a reference database), noting the uncertainties in these frequencies as estimates of the frequencies in the population.

   B. Present estimates of the relative frequency of a feature set in a relevant population based on a probabilistic model that relates the probabilities of combinations of features to the probabilities of individual features if the model is well grounded in theory and data.

   C. Present the classification from an automated system for making classifications if the sensitivity and specificity of the system have been established in relevant experiments. Explain these operating characteristics (or related quantities such as the conditional probabilities of incorrect classifications).

6. Forensic experts should confine themselves to speaking of the weight of the evidence (the support it lends to the parties' claims, e.g., "this is strongly indicative of identity – we expect to find it 10,000 times more often when it comes from the suspect than when it comes from a coincidentally matching person").

7. Forensic experts **should not** follow the current paradigm of opining on the claims themselves (e.g., "It's the defendant's fingerprint!")? Nor should the forensic expert use phrases such as "to a reasonable degree of scientific certainty" implying something statistical when the there is no statistical form to properly characterize the evidence.