THE ADMISSIBILITY OF "PROBABILITY EVIDENCE" IN CRIMINAL TRIALS—PART I

D. H. Kaye*

My last "Reflections" column talked about an employment discrimination case in which the Fourth Circuit underwent "hypergeometric confusion" in understanding a certain kind of probability (called a P-value) that can be helpful in evaluating the significance of statistical evidence. In this issue, I want to talk some more about the P-value, but in a very different context. The P-value, it seems, has insinuated itself in the criminal realm. Traveling incognito or under an alias, it has been used as a measure of the probative value of identification evidence—things like bloodstains, hair fibers, fingerprints and forged documents.

At first blush, it might seem that there is no place for probability calculations in the evaluation of this kind of nonquantitative evidence. Indeed, it was not so long ago that Marshall Houts observed:

> Unfortunately, it has become the vogue in recent years to bolster sagging cases by applying the law of probability in an attempt to strengthen weak or nonexistent evidence.
>
> I have discussed the matter with several mathematicians and they are unanimous in their opinions that the law of probability has absolutely no application to the forensic field. Its mathematical utility is founded on exact statistical research: and in no field of proof has this research progressed to a point sufficient to warrant the use of precise mathematical equations. I am further advised that this type of research required is not likely to come about within the foreseeable future.[1]

When Houts (who was, I am told, a trial lawyer in California) wrote these words thirty years ago, cases involving testimony giving the probabilities of

---

*D. H. Kaye is Professor of Law and Director, Center for the Study of Law, Science and Technology, Arizona State University, Tempe AZ 85287.

[1]M. Houts, From Evidence to Proof 132 (1956).

disputed or uncertain events usually were confined to specialized fields, such as probability sampling in trade name cases. In the criminal field the few instances of "probability evidence" were sufficiently outrageous to justify Houts's skepticism. Let's look at two of the more well-known cases.[2]

The most famous—or infamous—is probably the 1899 prosecution of Alfred Dreyfus. Dreyfus, a Jew who had risen to the rank of Captain in the French Army, was suspected of transmitting state secrets. Part of the evidence used to convict him of treason were letters that Dreyfus had written and that the prosecution said were cipher messages. To demonstrate that the messages were in fact written in some secret code, it was shown that in these documents the letters of the alphabet occurred in proportions that differed from typical French prose.

The defense argued that although the precise distribution of letters may have been unusual, some such departure from the average was not especially unlikely. The world-renowned French mathematician Henri Poincaré attempted to convince the hostile tribunal that among the many possible proportions in which the letters might appear, any particular set of proportions—even the most likely—is relatively improbable. By way of analogy, consider tossing a balanced coin one hundred times. The single most probable outcome is fifty heads and fifty tails, but the probability of this specific outcome is less than 1/10. Nevertheless, even the impassioned declaration from another mathematician—"Give me the works of Racine and I will show you that he, too, by your foolish tests is a traitor, for the works of Racine, like the letters of Dreyfus, do not show the most probable distribution"—fell on deaf ears. In proceedings that may have been tainted with anti-semitism, Dreyfus was convicted.

In 1915, not long after the Dreyfus affair, in a rather less celebrated New York case, another defendant was convicted partly on the strength of "probability evidence." The charge in *People v. Risley*[3] was fraud. The accused was an attorney. The fraud allegedly occurred when Risley was representing a corporate client in a civil matter. Risley could hardly be faulted for his devotion to his client's cause. Apparently, he had removed a document that had been placed in evidence in the civil case and typed in the words "the same" to make the meaning more favorable to his client. An expert on typewriters testified that as typed on the document, the six distinct letters in the words "the same" showed eleven specific peculiarities. For example, the "t" was not strictly vertical, but slanted, other letters were missing serifs, and so on. This typewriting expert reported that a typewriter taken from Risley's office produced characters with the same peculiarities. A second expert, described by the New York Court of Appeals as "a professor of mathematics in one of the universities in the state," testified that if one assumes that the probability that a typewriter

[2]For a thoughtful and sophisticated discussion of another early case, *see* Meier & Zabell, *Benjamin Pierce and the Howland Will*, 75 J. Am. Statistical A. 497 (1980).
[3]214 N.Y. 75, 108 N.E. 200 (1915).

would have a character with any given peculiarity is one-half, then the probability of drawing at random the exact set of peculiarities found in Risley's typewriter and on the altered document would be one over four billion. Notwithstanding a dissent that argued that the error, if any, in admitting this testimony was harmless, the Court of Appeals reversed Risley's conviction, observing that "[t]he statement of the witness was not based on actual observed data, but was simply speculative. . . ." The difficulty, in other words, lay in the mathematician's assumption that the probability of each and every peculiarity was one-half. Without data to support this premise the resulting figure of one out of four billion may well have been far from the true probability.

Dreyfus and *Risley* illustrate the kinds of elementary mistakes that can occur when probabilities become the subject of courtroom testimony. In *Risley*, the prosecution sought to demonstrate astronomical odds against the defendant's having a typewriter with the incriminating characteristics, unless of course it were the typewriter used to perpetrate the fraud. The number calculated, however, does not give this probability, and there is good reason to doubt that the computed probability pertains to anything in the case.

In *Dreyfus*, the problem is one of misinterpretation. The probability, even if correctly computed, of the letters in a document having the precise distribution found in Dreyfus's correspondence is not of much interest. What would be more useful is the probability of ordinary correspondence having a distribution as aberrational or even more aberrational than Dreyfus's supposed secret messages. This latter probability would be a P-value—the probability of extreme data arising under the null hypothesis.

But *Dreyfus* and *Risley* are old cases, separated from today's decisions by the better part of a century. They shed no light on what should happen when a party presents a well-founded P-value that bears on a material issue. Despite these cases—and more recent incarnations of them—Houts was wrong. The mathematics of probability has important applications to forensic proof, and the future that Houts and the mathematicians he consulted dismissed as unforeseeable is the present. In trying constitutional and statutory claims of discrimination, many attorneys are turning to increasingly sophisticated statistical models and explicit hypothesis testing.[4] Medical experts in many jurisdictions testify to something they call (incorrectly) the "probability of paternity."[5] Econometric analyses are appearing in antitrust and other commercial cases.[6] In criminal cases involving microanalysis of hair fibers, questioned documents, certain techniques in analytical chemistry and blood grouping, testi-

---

[4]*E.g.*, STATISTICAL METHODS IN DISCRIMINATION LITIGATION (D. Kaye & M. Aickin, eds. 1986).

[5]*E.g.*, Ellman & Kaye, *Probabilities and Proof: Can HLA and Blood Group Testing Prove Paternity?* 54 N.Y.U.L. REV. 1131 (1979).

[6]*E.g.*, Rubinfeld & Steiner, *Quantitative Methods in Antitrust Litigation*, 48 LAW & CONTEMP. PROBS. 69 (1983).

mony about probabilities, while by no means ubiquitous, is becoming increasingly common.

A handful of courts have set sail squarely against these developments, holding that numerical statements of probabilities or closely related quantities are never admissible in criminal cases.[7] Commenting on this slim line of cases in a talk last month, one psychologist remarked that "[f]rom a scientific viewpoint this is, of course, highly irrational. . . ." On the other hand, nobody wants a wholesale acceptance of mathematical methods that are ill-adapted to our system for resolving criminal complaints. In the next installment of this column, I shall focus on the modern cases in which probabilities or comparable numbers are introduced to prove identity—what criminalists call "individualization." I hope to demonstrate the plausibility of a middle course between Procrustean exclusion of "probability evidence" and uncritical acceptance of "numerobabble."

[7] *See* Minnesota v. Boyd, 331 N.W. 2d 480 (Minn. 1983).