Remarks on Probability in Law: Mostly, a Casenote and a Book Review

Robert Birmingham
University of Connecticut School of Law

Follow this and additional works at: http://digitalcommons.uconn.edu/law_papers

Recommended Citation
REMARKS ON 'PROBABILITY' IN LAW: MOSTLY, A CASENOTE AND A BOOK REVIEW

Robert Birmingham*

Professor Posner's theme in his Article on the right to privacy, as in his comprehensive studies of law, is that courts and lawyers can use economics to understand and improve the process and the results of legal decisionmaking. In his indirect response to Posner, Professor Birmingham comments on other attempts to inject nonlegal terms into legal analysis.

Kyburg, beginning a paper: "I wish to argue here that 'probability' occurs but rarely and peripherally in scientific theories." But 'probability', the word, occurs often in legal theories, maybe most often when these theories are believed to be most scientific. In particular, the literature of law and economics is full of it.

Three observations: (1) philosophers and others have recently done much work on the foundations of probability, so that these are confused; (2) 'probability' is interesting etymologically, in a way related to law; (3) almost always at law, 'probability' is used without regard to, or in ignorance of, (1) and (2).

Various problems, perhaps also opportunities, come from this; it is these I talk about. I do not present a sustained argument, but separate discussions, loosely connected in that each has to do with 'probability' as it is related to law. Much may be done without being very technical.

I.

The Dean of the Faculty of Law of the University of Flores . . . took up the most advanced position. "The fundamental spirit of democracy," he told a meeting of the Faculty of Law a few days after the revolution, "proclaims that it is better that a hundred desperate criminals escape than that one innocent man be lodged in jail."

—John K. Galbraith²

---

* Professor of Law, University of Connecticut. A.B., J.D., Ph.D., University of Pittsburgh; LL.M., Harvard Law School.


² J. Galbraith, The Triumph 126 (1968).
First, a pretty use of 'probability', introduced into law, so far as I know, by Kaplan. A person accused of a crime did it or not and may be found to have done it or not: four outcomes, each with a satisfaction, say to society. If the satisfaction, or lack of it, is with the decision, it maybe only matters if one gets it wrong.

$$\begin{array}{c|cc}
\text{Actuality} & \text{innocent} & \text{guilty} \\
\hline
\text{innocent} & 0 & s_{12} \\
\text{guilty} & s_{21} & 0 \\
\end{array}$$

In general, $s_{12} \neq s_{21}$: it is worse, say, to punish an innocent person than to let a guilty person go free. Be this as it may, if $p$ is the probability a person is guilty, it is a matter of indifference, in terms of satisfaction, whether the person is punished if

$$ps_{12} = (1 - p)s_{21}.$$ 

For each verdict, one multiplies the loss from mistakenly reaching it, $s_{12}$ or $s_{21}$, by the probability of doing so, $p$ or $1 - p$: thus expected (dis)satisfaction. If the products are equal, how one decides does not matter.

If one determines the $p$ at which the products are equal, one determines the ratio of $s_{12}$ to $s_{21}$ and finds out something about society, or persons in it. Happily, a study does this, giving, for jurors, mean probabilities from .74 for petty larceny to .92 for manslaughter; murder comes out .86. So much accords well with earlier estimates of the ratio of $s_{12}$ to $s_{21}$, mostly rhetorical, parodied by the epigraph.

---

4 As I explain better in part III, the probability involved is subjective: the defendant did it or did not, so objectively the probability is 1 or 0; but a juror is not given this.
Some subjectivists have asked, 'if the coin is flipped but the outcome is hidden from you, (a) wouldn't you still bet as though heads is a 50-50 chance, and, if so, (b) would you still insist on using the familiar (non-Bayesian frequentist) language that, once the coin is flipped, the probability is no longer ½ that the coin may come up "heads", but is either 0 or 1 (you know not which) that it now is heads?' My answer to both questions is yes.
REMARKS ON ‘PROBABILITY’

Hale: 5-1
Blackstone: 10-1
Fortescue: 20-1

In all this there is, of course, the presumption that some innocent persons are punished. The prospect was met with equanimity by Paley, who argued that “courts of justice should not be deterred” by “the mere possibility of confounding the innocent with the guilty” but “ought rather to reflect, that he who falls by a mistaken sentence, may be considered as falling for his country.” It wasn’t his ox.

II.

Suppose one takes a well-shuffled pack of cards, it is quite likely or not unlikely that the top card will prove to be a diamond: the odds are only three to one against; but most people would not say that it is quite likely to be the nine of diamonds for the odds are then fifty-one to one against. On the other hand I think that most people would say that there is a serious possibility or a real danger of its being turned up first and, of course, it is on the cards.

—Lord Reid

A case, done badly: The Heron II (Koufos v. C. Czarnikow, Ltd.). The plaintiffs chartered the defendant’s ship, the Heron II, to carry 3000 tons of Hungarian sugar from Constanza, on the Black Sea, to Basrah. (One has to get into the spirit of the thing: Basrah, at the head of the Persian Gulf, was founded as a garrison town of the conquering Arabs in 638; that is, it’s not, say, Cleveland.) The voyage, which should have taken twenty days, took eleven more, at least nine of them due to deviations in breach of contract. (The ship left Constanza 1 November 1960, arrived at Basrah 2 December 1960; but by contract it might have arrived for loading as late as 10 November.) The plaintiffs intended to sell the sugar on arrival of the ship at Basrah, and did so. The defendants did not know the plaintiffs intended to do this but knew there was a market for sugar

---

9 Id. at 686.
in Basrah and must have known the plaintiffs were not going to eat the sugar themselves. During the days the ship was delayed, the price of sugar in Basrah fell from £32 10s. per ton to £31 2s. 9d. per ton, presumably at least in part because another ship arrived there, as scheduled, delivering 8000 tons of Formosan sugar. (Gilmore says the market “collapsed”; he overstates the change.)

The defendant admitted liability of £172, mostly interest at six percent on the value of the sugar for nine days. The plaintiffs sought loss-of-market damages and recovered them. In addition to the £172, the umpire awarded more than £4000, because of the decline in the price of sugar during the delay. The trial court set aside the second part of the award, but the Court of Appeals reinstated it; the House of Lords affirmed the decision of the Court of Appeals.

The House of Lords, of course, decided the case under the rule of Hadley v. Baxendale:

Where two parties have made a contract which one of them has broken, the damages which the other party ought to receive in respect of such breach of contract should be such as may fairly and reasonably be considered either arising naturally, i.e., according to the usual course of things, from such breach of contract itself, or such as may reasonably be supposed to have been in the contemplation of both parties, at the time they made the contract, as the probable result of the breach of it.\[13\]

With some variation among the judges, the reasoning went this way: The defendant, if he had thought about it, must have realized (1) that it was, say, not unlikely that the sugar would be sold in the market on arrival, and (2) that market prices change. But, in spite of a tendency of the price of sugar at Basrah “to decline during October and November to a low point in about December,” the defendant might properly have taken there to be “an even chance that the fluctuation would be downwards.”

As the judges understood it, the issue, under Hadley v. Baxendale, was how probable what happened had to be. Their judgments about past cases, very roughly:

\[14\] [1867] 3 All E.R. at 697.
\[15\] Id. at 699.
REMARKS ON 'PROBABILITY'

Asquith, L.J., in Victoria Laundry (Windsor), Ltd. v. Newman Industries, Ltd.\textsuperscript{18} (mere) possibility on the cards serious possibility real danger grave risk liable to result

Alderson, B., in Hadley v. Baxendale not unlikely likely quite likely even chance more than even chance odds on chance

Viscount Dunedin, in Re R. & H. Hall, Ltd., and W.H. Pim (Junior) & Co.'s Arbitration\textsuperscript{17} reasonably certain virtually certain certain inevitable

Mellish, L.J., in The Parana\textsuperscript{16} reasonably certain virtually certain certain inevitable

\textbf{Hadley}, of course, was rightly decided; the problem was to interpret it. (A particular problem is a willingness to suppose, say, that something is reasonably certain if it is reasonably certain that it is reasonably certain—especially, that it is not unlikely if there is an even chance that it is not unlikely. See part IV.)

The reasoning of the judges, burdened by introspection about ordinary language, as by Lord Reid, above,\textsuperscript{10} mostly displays the disadvantages of a classical education as to this sort of thing. The opinions take up better than thirty pages; one waits, desperately, for somebody to say ‘about .1’ and be done with it. (As Dawson and Harvey see it, the opinions are “a stunning display of prolixity”;\textsuperscript{20} they disliked reading them and relieve the student from doing so.) This use of ordinary language is bad by itself, but there is a more substantial difficulty.

In 1931, somebody named Heisel, who disliked irrational numbers, tried to prove \(\pi\) is a rational number, \(256/81\), like this:\textsuperscript{21} Using \(256/81\) for \(\pi\), he calculated circumferences and areas of circles with diameters from one to ten. In each instance, the ratio of the area to the circumference was \(r/2\), as it should be. \textit{Q.E.D.}

\textsuperscript{16} [1949] 2 K.B. 528, 540.
\textsuperscript{17} [1928] All E.R. 763, 766, 767 (H.L.).
\textsuperscript{18} 2 P.D. 118, 123 (C.A. 1877).
\textsuperscript{19} \textit{See} text accompanying note 8 \textit{supra}.
\textsuperscript{20} J. Dawson & W. Harvey, \textit{Cases and Comment on Contracts} 70 (3d ed. 1977).
\textsuperscript{21} C. Heisel, \textit{Behold! The Grand Problem No Longer Unsolved: The Circle Squared Beyond Reputation} (1931).
Beckmann, who reports this, remarks, "he would have obtained the same consistency had he set $\pi$ equal to the birthdate of his grandmother."\textsuperscript{2} The judges in *The Heron II* do something like what Heisel did.

In cases like *The Heron II*, the defendant has breached a contract with the plaintiff, with consequent loss to the latter. As the judges see it, the defendant is liable for the loss if the event that caused it is *not unlikely*. Reconstructing a bit, one gets a function from events, understood to be particulars, to real numbers from 0 to 1, giving the probabilities of these events. If the number that is the value of this function for the appropriate event as argument of it is, for instance, above .1, the defendant is liable for the loss; otherwise he is not.

Unhappily, events, being particulars, do not have probabilities or do not have probabilities that one can use this way. The particular event that caused the loss to the plaintiffs in *The Heron II* was the arrival of the ship with the Formosan sugar, the sale of this sugar in the market at Basrah, etc. In a sense, if everything that happens is determined to happen, the probability of this event is 1. But it is not this sense that matters. What matters is the probability of this event in the light of what the defendant knew or should have known: *this* probability is 0. The first mate of the ship with the Formosan sugar was, perhaps, Lars, a Norwegian. If the first mate had been Ch'ien-fu, a Chinese, who did not exist, but might have existed, the event would have been different. So, too, would it have been different if the ship had arrived $\pi$ seconds later than it did. There are uncountably many possible first mates and times of arrival, each equally likely, so far as the defendant knew or should have known. Again, for the defendant, the probability of the particular event happening that did happen is 0.\textsuperscript{23}

The probabilities that matter belong not to events as such, but to *events under descriptions*.

Our ordinary talk of events, of causes and effects, requires constant use of the idea of different descriptions of the same event. When it is pointed out that striking the match was not suffi-

\textsuperscript{2} P. Beckmann, *A History of $\pi$* (Pi) 175 (2d ed. 1971).

\textsuperscript{23} Illustration: if one shoots a machine gun accurately at a target of area 1 and idealizes a bit, so that bullets are point-particles, a region of area 0 is hit each time, that is, the region of all and only the points hit in the long run; the rest of the target, still with area 1, is not hit at all. van Fraassen, *Relative Frequencies*, 34 *Synthese* 133, 135 (1977).
REMARKS ON ‘PROBABILITY’

So what did happen, described as the arrival of a ship, had a particular probability; described as the arrival of a ship carrying sugar, or carrying 8000 or more tons of sugar, it had different, lower probabilities.

The catch is, of course, that there is no correct description. In The Heron II, Lord Morris understood the description that counted to be ‘a loss by reason of a fall in the market price of sugar at Basrah’ and believed the event, thus described, “was liable to result or at least not unlikely to result.” But a change in the market price was virtually certain, and a fall in the market price of four percent or more unlikely but a serious possibility. And maybe it should matter whether the market price fell because a ship came in, or because the wife of the trader Abû-Nuwâs, until then even-tempered, threw a pot at him, so that he decided to sell out and go to Kuwait. The court spends thirty pages deciding which probability is important, but presupposes the description, equally determinative, under which the event is given a probability to be compared with it. Its presupposition, itself arbitrary, pretty much decides the case.

(As Davidson does, I take events to be particulars. If I take them to be not particulars, but properties, perhaps properties of points in space-time, the problem does not go away but is described differently: the judges must decide among properties, that is, events instead of descriptions of them. They still don’t do it.)

III.

The Court: Now, if that is the case, you must—before you hear any evidence at all—you must start on the theory and believe that this man is innocent. But as soon as they produce proof which satisfies you beyond a reasonable doubt that he is guilty, then you can feel otherwise. Do you understand that?

The Juror: Yes.

—Voir dire conducted by Charles R. Garry in People v. Newton

26 Davidson, supra note 24.
I go back to the discussion of determining guilt or innocence in part I. In a sense, again, the accused either did it or did not, so the (objective) probability he did it is either 1 or 0. But the probability that matters is subjective: if a juror believes, for instance, that the probability the accused did it is .9 or better, he votes to convict him. I asked above what the juror does, given the probability the accused did it; I ask here how the juror arrives at this probability.

In part, it is easy. A form of Bayes's theorem, or an application of it, is

\[
P (\text{guilty} \mid E) = \frac{P (E \mid \text{guilty}) \cdot P (\text{guilty})}{P (E)}
\]

Read: 'The probability a person is guilty, given evidence \(E\), is the prior probability he is guilty, multiplied by the probability of \(E\) given he is guilty, divided by the prior probability of \(E\).' The idea is that the juror starts with a probability of guilt, which is changed as he gets evidence. Lempert \(^{29}\) discusses Bayes's theorem in an interesting way, but I give my own examples.

Suppose a person accused of being a witch has "a frightened look." \(^{31}\) If the prior probability of being a witch is \(p\), but everybody accused of this has, not unreasonably, a frightened look, the probability of her being a witch is still \(p\),

\[
P (\text{guilty} \mid E) = \frac{1}{1} \cdot p = p
\]

and one hasn’t found out anything. (The law did not look at it this way. As a handbook for inquisitors put it, that a person accused of being a witch has "a frightened look" is "a light indication" of guilt, by itself "sufficient warrant for proceeding to the torture." \(^{32}\)

Again, suppose a prosecutor produces "two qualified witnesses to testify that they heard the sounds of a violent struggle in a house, that they saw the accused emerge from the house with a blood-stained knife in his hand, and that the house was empty save for the body of the victim." \(^{33}\) If the prior probability the accused did it is \(p\), and the probability of the evidence given he did it is 1, but the evidence is otherwise not probable, perhaps .01, then the probability he did it goes up a hundredfold:


\(^{31}\) H. Bouquet, An Examen of Witches 223 (1929).

\(^{32}\) Id.

\[ P(\text{guilty} \mid E) = \frac{1}{.01} \quad p = 100p. \]

(But, in Islamic law, “Shari’a doctrine forbids the judge to draw from this evidence the conclusion that the accused was the killer,” this evidence being “what is termed ‘suspicion.’”31)

The epigraph is an interjection by the judge in the questioning by Garry of a prospective juror.

Q. Well, my question is: As you sit there right now, do you believe that Huey Newton shot and killed, stabbed, whatever it was, Officer Frey?

A. I don’t know whether he shot him or not. That I can’t say.

The Court: Mr. Strauss, you see, under our law there is a presumption of innocence to start with. When you start the case the defendant is presumed to be innocent, and it is up to the People, the prosecution, to prove to you beyond a reasonable doubt that the defendant is guilty. Do you understand that?

The Juror: Yes.

The Court: So, now, not having heard any evidence, you must start with a presumption of innocence. Do you know what I mean by presumption? You must say, “As far as I know the man is innocent.” Do you understand that?

The Juror: Yes.35

It ended badly.

Mr. Garry: As Huey Newton sits here next to me now, in your opinion is he absolutely innocent?

A. Yes.

Q. But you don’t believe it, do you?

A. No.

The Court: Challenge is allowed.36

In light of Bayes’s theorem, one obvious way to interpret the presumption of innocence is to take it to require that the prior probability a person is guilty, the probability a juror starts with, be 0. The point is, this is not possible. Then

---

31 Id.
32 Id. at 94.
33 MINIMIZING RACISM IN JURY TRIALS, supra note 28, at 90-91.
34 Id. at 94.
with $p = 0$. But now, no matter what the evidence, there is no way to get the posterior probability a person is guilty, the probability a juror ends with, to be anything other than 0, because 0 multiplied by any number is 0: "If the prisoner is to have a greater than zero probability of guilt on the evidence, the calculus of change requires that his prior probability of guilt must also be greater than zero."37

IV.

For example, the probability of a number's being prime, if greater than 10 and less than 20 may be said—informally—to be .5, because out of eight such numbers just four are primes.

—L. Jonathan Cohen48

A short answer is, of course, one need not be a Bayesian about this;39 but here a book by Cohen comes in: The Probable and the Provable. "What is now the standard theory of probability," says van Fraassen, "is very simple. A probability measure is a map of a Borel field of sets into the real number interval [0, 1], with value 1 for the largest set, and this map is countably additive."40 That is, one has a set of individuals, assigned weights. The probability of one of them is its weight; the probability of a set of them is the sum of the weights of the individuals in it. Cohen, although he gets off to a bad start, as by the epigraph, renounces this approach, which he disparages as ‘mathematicist’.

It does not, he believes, apply at law, for six reasons.41 (1) In civil
cases, a plaintiff has to prove all aspects of his claim, but not with probabilities sufficiently high that the product of them exceeds .5. (2) Inferences on inferences are said (Cohen cites Wigmore) to be bad things, so that a first inference must be proved beyond a reasonable doubt. (3) Sometimes, as when one is run over by a bus but is not able to identify the bus line, a probability by itself will not do. (4) In criminal cases, juries do not decide whether there is a reasonable doubt by probabilities, as in part III, but by reasons for doubting (Cohen's example: the shifty demeanor of a witness). (5) More generally, no model of the mathematicist, e.g., the betting model, seems appropriate to reasoning in law, even as aspiration; the mathematician is obliged to come up with one. (6) As above, part III.

What Cohen is selling is something he calls 'inductive probability', which goes like this: Suppose a hypothesis, 'All ravens are black,' and n tests $t_i$ of it,

\[
t_1 = \text{observing a raven} \\
 t_2 = \text{observing a raven of each sex} \\
 t_3 = \text{observing a raven of each sex in each season}
\]

and so on. If the hypothesis passes $t_i$, it is supported to degree $i/n$; the hypothesis 'Not all ravens are black', is unsupported.

Applied to law, the approach requires a plaintiff to support his position, that is, show that his hypothesis, $p$, passes a test; the probability of it is then positive; the probability of the defendant's

---

*b Cohen gets into trouble here, as he does in the epigraph.

[Suppose that there is expert evidence asserting the presence of just three mutually independent peculiarities in the machine on which the letter was typed, and of just the same three in A's machine. Suppose the expert evidence is that each of these peculiarities is normally found, on average, in only one out of ten machines. Then, if this evidence is correct, the mathematical probability of a typewriter's having all three peculiarities will be .001; and the mathematical probability that any randomly selected pair of typewriters have, independently of one another, all three peculiarities will be one in a million. It is tempting thence to infer a .999,999 mathematical probability that the letter's having been typed on a machine with those peculiarities is dependent on the fact that A has such a machine, and therefore apparently a .999,999 mathematical probability that A typed the letter.

L. COHEN, supra note 38, at 84. It is not tempting at all. Consider the airplane passenger who brought a bomb because he was pretty sure it would not go off and reasoned that the probability of one bomb on a plane being $p$, the probability of two is $p \times p$. This from my student, Gil Walker, who helped elsewhere, too. See also People v. Collins, 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968) (en banc).

hypothesis, not-\(p\), is 0. The defendant must then show that his hypothesis passes a higher test, so that the probability of his hypothesis is positive and the probability of the plaintiff’s hypothesis 0; or the defendant must at least offer evidence so that the probabilities of both hypotheses are 0. Plainly, if probabilities are obtained this way, it is not possible to decide what to do, as in part I, by multiplying probabilities and satisfactions.

But if a probability is not to be used in a calculus of this kind, it is not clear what good it is.

Mathematicist: So set the inductive probability of the unsupported hypothesis, say not-\(p\), at 0; but let the inductive probability of \(p\) be perhaps the mathematical probability of \(p\) minus the mathematical probability of not-\(p\), so that inductive probability may be translated into mathematical probability.

Inductivist: That’s just the point. It can’t be.

Mathematicist: But then why should jurors, say, pay any attention to it? They must, at least implicitly, decide as in part I, by expected satisfaction. “In the words of Edmund Morgan, ‘How else can they think?’—at least think rationally.”\(^4^5\) If the cash value of inductive probability is not mathematical probability, jurors might as well examine the entrails of chickens.

Inductivist: I don’t know.

The dialogue might end this way. If it does not, perhaps it goes on as follows.

Inductivist: My approach does not maximize (expected) satisfaction as in part I. But it has institutional advantages, e.g., everybody believes it to be fair: \(^4^6\) satisfaction may nevertheless be maximized by it.

Mathematicist: I am reminded of the utilitarian dilemma by which a sheriff is required to frame an innocent person to stop lynchings that will take place if a guilty person is not found or believed to be found.\(^4^7\)

\(^{4^5}\) Kaplan, supra note 3, at 1070. Kaplan adds: “Despite the urging, indeed importuning, of several delegations of law review editors, the author has been unable to discover when or where Professor Morgan said this. In any event, the reader may rest assured that he did so.” Id. at 1070 n.6.


\(^{4^7}\) See generally McCloskey, An Examination of Restricted Utilitarianism, 66 Phil. Rev. 466 (1957).
Or maybe satisfaction does not matter; but if it does not, again, it is not evident what does.

But the mathematicist’s position, unsupplemented, won’t do either. The trouble begins with the juror of part III, who, one recalls, refused to presume Newton innocent: if innocent, what was he doing there? Suppose a prosecuting attorney trying a case of assault argues this way:

The accused is being tried for assault. Jurors convict for assault if the probability the accused did it is at least .75. The probability an accused did it, as determined by judges, is .85.

Therefore: one ought to convict.

And stops. The idea is, if what he says is true, and if the accused does not add anything, society is better off if the accused is convicted. But the law does not permit this.

The problem is not only a legal one. One supposes of a coin one knows nothing about that \( P(\text{heads}) = P(\text{tails}) = .5 \), but does so uneasily. If one tosses it one hundred times and gets fifty heads and fifty tails, one supposes the same thing, more easily.

In 1876, Stanley, back in Africa after finding Livingston, had with Frank to decide whether to go down the Zambezi (south) or, more dangerously, because it had not been done before, down the Congo (north). “Finally, Frank came up with an idea: ‘I say, sir, let us toss up: best two out of three to decide it. . . . Heads for the north and the Lualaba; tails for the south and Katanga.’” They did so, with an improbable result. “Stanley gave Frank a rupee coin. The youth flipped it; he flipped it in fact six times and six times it came up tails for the south and Katanga.” They went down the Congo anyway, and Frank drowned in it.

How would six successive tails have changed the probabilities associated with the coin? If they occurred on tosses 1-6, there being only six tosses, perhaps \( P(\text{heads}) = .2; P(\text{tails}) = .8 \). But if they occurred on tosses 101-106, with tosses 1-100 as above, perhaps \( P(\text{heads}) = P(\text{tails}) = .5 \) as before.

The difference is in the resiliency of the probabilities. Again

---

4 Simon & Mahan, supra note 5, at 328.
6 It is, but it doesn't quite fit. Judges require a probability of guilt of .88 to convict for assault. Simon & Mahan, supra note 5, at 328. Jurors believe .69 of persons tried are guilty. Kalven, supra note 49, at 1064.
8 Id. See also T. STOPPARD, ROSENKRANTZ AND GUILDENSTERN ARE DEAD (1968).
somewhat ideally, all of us have belief states, functions from sentences of our language to probabilities of the truth of them. These states change, as when we find out a sentence, that is, the truth of it, has probability 1 and adjust probabilities of other sentences in light of this. The resiliency of the probability of a sentence with respect to a set of sentences is defined as 1 minus the greatest difference between the probability of this sentence to begin with and its probabilities obtained by alternately taking each of the sentences in the set to have probability 1. (Then the resiliencies of both $P$ (heads) and $P$ (tails) are .7 and 1 for the sets of sentences \{‘Tosses 1-6 are tails.’\} and \{‘Tosses 101-106 are tails.’\} respectively.)\(^{13}\)

What is wrong with the way the prosecuting attorney argues above, and also, I believe, with the case against the bus company,\(^4\) is that the probabilities of guilt and responsibility, respectively, even if high, are not resilient enough: they change very substantially if the belief states with which they are associated are changed by taking, as is plausible, obviously pertinent sentences, for instance, ‘The person assaulted says, “The accused didn’t do it.”’, to have probability 1. I do not know how high resiliency ought to be, or what sentences ought to be in the set by which one determines resiliency. The point, though, is that Cohen discards where he should supplement. Mathematical probabilities matter; it is just that they are not all that matter.

V.

Sir, the law is as I say it is, and so it has been laid down ever since the law began; and we have several set forms which are held as law, and so held and used for good reason, though we cannot at present remember that reason.

—Chief Justice Fortescue\(^5\)

The strategy Cohen adopts is that of showing that proof at law is by inductive probability. (For instance, as to the first of the six reasons that mathematical probability does not apply at law,\(^5\) i.e., that a plaintiff has to prove all aspects of his case but just barely, he says, if every part of a plaintiff’s case has a positive inductive probability, the conjunction of the parts has too; but the denial of

\(^4\) See text accompanying note 42 supra.
\(^5\) Y.B. 36 Hen. 6, 25b-26 (1458).
\(^{14}\) See text accompanying note 41 supra.
the conjunction of the parts has probability 0.) As I see it, on the other hand, a theory of probability is not just descriptive, the point of it is not merely to model the way one reasons at law; it is norma-
tive and, if necessary, revisionary.

As a place to start, procedures of proof at law were once different from what they are now.

The German laws refer to cases in which a woman might demand justice of a man personally in the lists, and not only are instances on record in which this was done, as in a case at Berne in 1228, in which the woman was the victor, but it was of sufficiently frequent occurrence to have an established mode of procedure, which is preserved to us in all its details by illu-
minated MSS. of the period. The chances between such une-
qual adversaries were adjusted by placing the man up to the navel in a pit three feet wide, tying his left hand behind his back, and arming him only with a club, while his fair opponent had the free use of her limbs and was furnished with a stone as large as the fist, or weighing from one to five pounds, fast-
tened in a piece of stuff. A curious regulation provided the man with three clubs. If in delivering a blow he touched the earth with hand or arm he forfeited one of the clubs; if this happened thrice his last weapon was gone, he was adjudged defeated, and the woman could order his execution. On the other hand, the woman was similarly furnished with three weapons. If she struck the man while he was disarmed she forfeited one, and with the loss of the third she was at his mercy, and was liable to be buried alive. According to the customs of Freisingen, these combats were reserved for accusations of rape. If the man was vanquished, he was beheaded; if the woman, she only lost a hand, for the reason that the chances of the fight were against her.57

In part, this has less to do with proof than with efforts to restrict resort to force, independent of a concern to adjudicate. But wager of battle was a way to get at truth, too: champions of equal prowess were chosen so that God’s testimony might more easily be re-
ceived;58 by his loss the loser became “a convicted perjuror.”59 (This aspect of truth-seeking is unadulterated in the ordeal.)

57 Lea, The Wager of Battle, in Law and Warfare 233, 244 (P. Bohannan ed. 1967).
It is not as though while this was going on there was a system of probabilistic adjudication, as in part I, that courts were conscious of but chose not to use. There simply was no concept of probability in our sense before 1654. The word 'probability' was used before then with a different meaning, now dropped. In 363, the Emperor Julian fought his way into Persia and died there at thirty-two, fighting without armor. His successor, Jovian, less effectual, his army ill-supplied, by treaty gave up five provinces to get out. Gibbon remarks: "According to Rufinus, an immediate supply of provisions was stipulated by the treaty, and Theodoret affirmsthat the obligation was faithfully discharged by the Persians. Such a fact is probable, but undoubtedly false." By this use something can be both probable and false. "The old medieval probability was a matter of opinion. An opinion was probable if it was approved by ancient authority, or at least was well testified to."\(^6\)

Thesis. The job of the jury has always been to establish probability, but it has done this in two radically disparate ways. (1) By making probable what it determines happened. Pollock and Maitland: "That in old times 'the jurors were the witnesses'—this doctrine has in our own days become a commonplace."\(^2\) The casualness about how jurors went about being witnesses suggests that the purpose of the jury was less to determine what happened than to make what it determined happened probable in the old (testimonial) sense. It still does—I suggest, vestigially. Others, for instance Tribe, if I understand him, believe it ought to go on doing so and not do much else.\(^3\)

(2) By establishing, in a different sense, that of 'finding out', what probably happened. Langbein: "The English substitute for the judgment of God was the petty jury, an institution that retained something of the 'inscrutability' of the ordeals"; he criticizes it as "so crude that torture was unnecessary."\(^4\) As I understand him, he takes the purpose of the jury to be to get at probability in the new

\(^{10}\) E. Gibbon, The History of the Decline and Fall of the Roman Empire 835 n.116 (Modern Library n.d.) (1st ed. 1776-1788). Cicero has: "Atque illud est probabilius, neque tamen verum, quod Socrates dicere solebat, omnes in eo, quod scirent, satis esse eloquentes." De Oratore, in 3 Cicero I.xiv.63. This gets translated: "What Socrates used to say, namely, that all men were sufficiently eloquent in the field they knew, has some probability, but no truth." The Socratic Enigma 28 (H. Spiegelberg ed. 1964). Spiegelberg doesn't give the date of the translation.

\(^{11}\) I. Hacking, The Emergence of Probability 43-44 (1975).

\(^{12}\) 2 F. Pollock & F. Maitland, supra note 69, at 622.

\(^{13}\) Tribe, supra note 46.

\(^{14}\) J. Langbein, Torture and the Law of Proof 77 (1977) (quoting T. Plucknett, Edward I and Criminal Law 75 (1960)).
(mathematical) sense and believes, justifiably, that in the beginning it did not do this very well. I think it was not then supposed to, but is now.

(1) and (2) are related like this. About 1660, "[a] new kind of testimony was accepted: the testimony of nature which, like any authority, was to be read." For the law, (1) and (2) are related partly by pun. If a jury reasons, it ought to do it right, i.e., by mathematical probabilities. But its role in (1) is mostly ritualistic. As to Cohen, the concept of inductive probability, I think, simply conflates (1) and (2), doing badly by both.

---

68 I. Hacking, supra note 61, at 44.