

Probability, policy and the problem of reference class

By Robert J. Rhee*

Associate Professor of Law, University of Maryland School of Law

The mathematical concept of probability is relevant in evidentiary assessment. The indeterminacy of fact-finding and legal assessment, unique products of human intuition and intelligence, is well known. The application of law is sometimes unpredictable, and thus the substitution of an apparently formal methodology for a subjective process is alluring. But many scholars, and more importantly courts, shun an overtly mathematised process.¹ In most situations, probability as a mechanical output, while relevant, is neither dispositive nor particularly weighty. Why is it so? And why should it be so? There are many reasons.² This short article focuses on the problem of reference class as it relates to probability and weight of evidence, and policies underlying legal rules.

Statistical probability requires a reference class from which a proportion is derived.³ In simple matters, of coin flips or sex of newborns, the appropriate reference class is trivially obvious. However, as Ronald Allen and Michael Pardo showed, evidentiary inquiries are seldom defined in such narrow symmetric parameters.⁴ Probability assessments change with the reference class. If a proposition is subject to proportional comparison against two or more different

* Email: rrhee@law.umaryland.edu. I thank Brad Borden, Alex Glashauser and Bill Merkel for their comments. I thank Ronald Allen for his generous invitation to participate in this symposium.

1 See e.g. *People v Collins*, 438 P 2d 33 (1968); Laurence H. Tribe, 'Trial by Mathematics: Precision and Ritual in the Legal Process' (1971) 84 *Harvard Law Review* 1329.

2 See e.g. Ronald J. Allen and Michael S. Pardo, 'The Problematic Value of Mathematical Models of Evidence' (2007) 36 *Journal of Legal Studies* 107; Charles Nesson, 'The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts' (1985) 98 *Harvard Law Review* 1357; Tribe, above n. 1.

3 John Maynard Keynes, *A Treatise on Probability* (Macmillan: London, 1921) 94.

4 Allen and Pardo, above n. 2.

references, their selection is often a subjective process. Thus, the advantage of objectivity and methodological rigour is illusory.⁵

Consider Jonathan Cohen's famous paradox of the gatecrashers.⁶ The sole and undisputed evidence is that 501 rodeo attendees out of 1,000 did not pay, though their identities are unknown. Can the rodeo owner sue each of the 1,000 attendees on the theory that there is a .501 probability that they crashed the gate? No one, not even the mathematician, disputes that the plaintiff should lose the case. The proffered reasons have differed in shades: 'policy' should prohibit the plaintiff from receiving a windfall;⁷ verdicts should not be made on the basis of a probabilistic 'bet';⁸ liability assessment should require 'particular' evidence.⁹ These explanations correctly suggest the unreliability from the juridical perspective of naked statistical evidence.

Cohen's thought on his paradox is grounded in the problem of reference class. If it is true that there is only one reference—the class of rodeo attendees divisible into payers and gatecrashers, like the flip of a coin—there may be no reason why each of the 1,000 defendants should escape liability. To be sure, there would be an error with a windfall of 499 to the rodeo owner, but this error could be countenanced since otherwise there would be an error with a windfall of 501 to the gatecrashers. The error is directional, and it does not inevitably defeat the plaintiff's claim since no one guarantees an omniscient process. Rather, the complication in this calculus is the problem of reference class. The rodeo attendees, as Cohen noted, could be categorised into ex-boy-scouts and non-scouts, the former perhaps tending to be more honest. 'That is, if the court were told all the circumstances that were relevant to the statistical probability of the man's having bilked the rodeo organizers, the reference class for calculating this probability would be rather different'.¹⁰ The probability statement, while containing relevant information, cannot alone establish liability since the reference class cannot be limited to one.

The civil standard of proof supports an inductive assessment process. A common formulation of the preponderance of the evidence standard is that a proposition

5 Keynes, above n. 3 at 85.

6 L. Jonathan Cohen, *The Probable and the Provable* (Oxford University Press: New York, 1977) 74–81. See *Smith v Rapid Transit Inc.*, 58 NE 2d 754 (Mass 1945).

7 David Kaye, 'The Paradox of the Gatecrasher and Other Stories' [1979] *Arizona State Law Journal* 101, 104; Tribe, above n. 1 at 1349.

8 Nesson, above n. 2 at 1379.

9 Tribe, above n. 1 at 1346.

10 Cohen, above n. 6 at 78.

must be more likely than not, which is sometimes restated as greater than 50 per cent of the evidence.¹¹ This suggests probability as a numeric concept, thus the instinctive attraction to mathematisation. But the definition of preponderance refers not only to 'superiority in number or amount',¹² but also 'greater weight of the evidence'.¹³ John Maynard Keynes articulated the difference between weight and probability as this: 'The weight, to speak metaphorically, measures the *sum* of the favourable and unfavourable evidence, the probability measures the *difference*.'¹⁴ In most cases, when there is only a naked statistic, there is little weight. While the probability between propositions x and not- x stated as a ratio may lean greatly to one or the other given the available evidence, the possibility of error may nevertheless be great because the amount of evidence creates a sufficient number of possibilities, yet unknown, such that reasonable minds cannot be confident in the truth of this proposition despite its probabilistic superiority over its counterpart given the evidence and the case theory.¹⁵ Sufficiency of weight, then, refers to a threshold of rational acceptance given the evidence *and* the belief in the quantum of unknowns relevant to each party's theory of truth. In other words, the proposition must be probable 'in the sense that actual belief in its truth, derived from the evidence, exists',¹⁶ and not simply *a* probability given the evidence and the parameters of the chosen reference class. According to Cohen, then, inductive probability has the required property to account for weight, but mathematical probability does not.¹⁷

11 See e.g. *Fennell v Southern Maryland Hosp. Ctr. Inc.*, 580 A 2d 206, 214 (Md 1990); *Russell v Subbiah*, 500 NE 2d 138, 141 (Ill App Ct 1986).

12 *Oxford English Dictionary*, 2nd edn (Oxford University Press: New York, 1989) 1289.

13 *Black's Law Dictionary*, 7th edn (Thomson West: St Paul, 1999) 1201. See *Oxford English Dictionary*, above n. 12 at 1289 ('The fact of exceeding in weight; greater heaviness').

14 Keynes, above n. 3 at 77. Both Keynes and Cohen recognised the importance of weight in the assessment of a probability proposition. Keynes did not believe that the concept of evidentiary weight had 'practical significance' (ibid. at 76). Cohen attributed Keynes' ambivalence to a wrongly conceived concept of a 'mere auxiliary to the theory of mathematical probability'. Cohen, above n. 6 at 36. Keynes, however, did clarify that weight and probability 'are independent properties'. Keynes, above n. 3 at 76. In matters of evidence, this must be the case. For example, assume that the first five pieces of evidence favour one party. The probability must increase substantially to account for the difference in the available evidence, and yet one may have low confidence since the weight of the evidence may still be low.

15 See Allen and Pardo, above n. 2 at 134 (noting the role of confidence in evidentiary assessment); Robert J. Rhee, 'A Price Theory of Legal Bargaining: An Inquiry into the Selection of Settlement and Litigation under Uncertainty' (2006) 56 *Emory Law Journal* 619, 655 (noting the role of confidence in case assessment).

16 *Sargent v Massachusetts Accident Co.*, 29 NE 2d 825, 827 (Mass 1940). See also *Guenther v Armstrong Rubber Co.*, 406 F 2d 1315, 1318 (3d Cir. 1969).

17 Cohen, above n. 6 at 81.

It has long been recognised that mathematical analysis of probability requires ‘not a wide experience or detailed information, but a completeness of symmetry in the little information there might be’.¹⁸ In most cases where this symmetry is lacking, acceptance of naked probability would be an exercise in ‘balanced ignorance’.¹⁹ In the case of the gatecrasher paradox, for example, one must ignore the possibilities of other reference classes: perhaps ex-scouts are more honest, perhaps younger people are more prone to crash the gate, perhaps richer or religious people are more likely to pay the fare. Given the array of possibilities and probabilities, one cannot conclude that based on *this* probability, framed by this reference class, the rodeo organiser is entitled to judgment. And, of course, the simplest gesture or testimonial phrasing at trial can easily outweigh even a high probability proposition because there can never be a way to objectively model the unique human experiences that assess credibility or frame the unique reference from which one perceives and understands the world.²⁰

The problem of reference class is not limited to the field of evidence. Evidentiary assessment naturally flows into case assessment, and the same problem is confronted here. In any legal dispute, the decision to try or settle a case depends in part on each party’s probability assessment of the juridical outcome. An agreement as to the probable outcome should always result in settlement, but in a contested action few parties agree (thus the dispute). Like the mathematicians in the field of evidence, law and economics scholars in bargaining theory are drawn to the idea of objectivity. Conventional law and economic analysis assumes the ‘rational expectation’ model, i.e. ‘the subjective expectations correspond to the objective frequencies of the random event’.²¹ Settlement is said to be efficient because it reduces transaction cost. This normative idea holds only if one assumes the existence of an objective probability of case outcome; for efficient settlement to occur, then, the parties must discover this probability. But lurking beneath this elegantly simple analysis is again the problem of a reference class. Case assessment can be framed from the reference point of the judge, the court and forum, the attorneys, the parties (if repeat players), the type of action, the type of injury, the legal framework, and—not the least of which—the evidentiary assessment. Multiple probability inferences exist on multiple planes, and one cannot expect consistency. Thus, the rational expectation model assumes an

18 Keynes, above n. 3 at 85.

19 Ibid.

20 See Ronald J. Allen, ‘Factual Ambiguity and a Theory of Evidence’ (1994) 88 *Northwestern University Law Review* 604, 625–6; George Pólya, *Mathematics and Plausible Reasoning: Volume II Patterns of Plausible Inference* (Princeton University Press: Princeton, 1968) 110.

21 Robert Cooter, ‘The Cost of Coase’ (1982) 11 *Journal of Legal Studies* 1, 22. See e.g. George L. Priest and Benjamin Klein, ‘The Selection of Disputes for Litigation’ (1984) 13 *Journal of Legal Studies* 1, 7.

objective probability, but as a matter of positive theory the assumption is illusory for the existence of multiple reference classes requires a subjective choice.²²

The commonality in the debate is the concern of error (cost) mitigation. The mathematicist's core contention is that a formal algorithmic process can reduce errors. In this regard, there are two concepts of error: error in the individual case and systemic error. The juridical process is a search for a plausible truth of the past.²³ Yet objective probabilities refer 'only to classes and never to individuals, except in so far as they imply or presuppose that the individual in question is a member of some class'.²⁴ Here, then, is the crux of the matter. Absent special circumstances, individualised justice is incompatible with the application of aggregate experience to imply the occurrence of a specific event consistent with the class, particularly when there are multiple classes to choose from and the choice, as the term suggests, is subject to multivariate rationality.²⁵ In other words, the determinative inquiry becomes the instrumental search for the proper reference point rather than the discovery of act, causation and injury. Implicit in the juridical process is the policy that factual ambiguity is best resolved through an attention to operative facts and their credibility, which together should form the probability proposition, and only the concept of inductive probability can comfortably accommodate this idea. There is no reason to believe that many iterations of this inductive process produce systemic error.

Under special circumstances, however, the mitigation of individual and systemic error is not positively related. Tort law provides a bounty of examples. Consider the famous cases of *Summers v Tice*,²⁶ *Sindell v Abbott Laboratories*,²⁷ and *Haft v Lone Palm Hotel*.²⁸ Each case presents a peculiar set of facts, conducive to a statistical or quasi-statistical analysis, under which the ordinary fact-finding process will not support liability, and yet the accumulation of these outcomes produces systemic error. In *Summers*, the probability was precisely equal that one of two defendants

22 See Rhee, above n. 15 at 638–46.

23 See Allen, above n. 20. See also *People v Collins*, 438 P 2d 33, 40 (1968).

24 A. J. Ayer, *Probability and Evidence* (Columbia University Press: New York, 1972) 51. 'Thus, the judgment which assigns to members of the class of cigarette-smokers such and such probability of dying of lung cancer is, as I have said, statistical; but the judgment that such and such an individual smoker will probably die of lung cancer, if it is genuinely a judgment about this individual, and not just about the class of smokers to which he belongs, is a judgment of credibility': *ibid.* at 28.

25 See Allen and Pardo, above n. 2 at 130–5 (discussing *United States v Shonubi*, 103 F 3d 1085 (2d Cir. 1997)).

26 199 P 2d 1 (Cal 1948).

27 607 P 2d 924 (Cal 1980).

28 478 P 2d 465 (Cal 1970).

acting in concert shot the plaintiff. In *Sindell*, market share data provided a rational proxy for the probability that the defendant's drug harmed the plaintiff. In *Haft*, the probability was very low, albeit unquantifiable, that a sign warning that no lifeguard was on duty would have prevented the plaintiffs' drowning. Ironically, an adherence to an objective probability concept in these cases results in 'recurring misses' that create systemic error.²⁹ Under these circumstances, particular rules of law, such as burden shifting and proportional liability allocation, are used to overcome the problem. These rules do not conflict with the concept of particularised justice, since the circumstances cannot support such treatment in these classes of cases. Importantly, an inability to discover facts due to epistemological difficulties in recurring cases with similar fact patterns is distinguished from an absence of evidence arising from ordinary circumstantial difficulties. The former can lead to systemic error requiring normative correction, while the latter is simply the product of the burden of proof. Thus, the statistical concept of probability is sometimes considered in the rule-making process, but only if such consideration is at least not inconsistent with the dictates of particularised fact-finding.

In conclusion, some general observations follow. As Allen and Pardo posited, the debate over the use of mathematics is not one of exclusion, but rather one of appropriate weight. The reluctance to inject mathematical formalism into the fact-finding function is justified. A legal dispute is the search for a plausible understanding of the truth, and an overtly mathematised process runs the risk of substantially substituting particularised inquiry with aggregate assessment, a notion inimical to transactional causation between conduct and injury. In special circumstances when such inquiry cannot advance the goal of error minimisation, a rule of law consistent with a class proposition may advance the notion of statistical justice. But as a general matter, the juridical process should not confuse aggregate assessment with justice *inter partes*.

29 Saul Levmore, 'Probabilistic Recoveries, Restitution, and Recurring Wrongs' (1990) 19 *Journal of Legal Studies* 691.

Copyright of International Journal of Evidence & Proof is the property of Vathek Publishing Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.